

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 01:54:26 ; Search time 2.78648 Seconds
(without alignments)
912.596 Million cell updates/sec

Title: US-09-593-793A-113_COPY_367_375

Perfect score: 48

Sequence: 1 GATCLSHSV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	123	4 AAM01228	P5538 spl
2	48	100.0	123	4 AAU69873	Human pro
3	48	100.0	123	4 ABU71764	Prostate
4	48	100.0	123	5 ABB95333	Human P55
5	48	100.0	123	6 ABR54445	Prostate
6	48	100.0	123	7 ADB14156	Human pro
7	48	100.0	231	3 AAY54369	Amino aci
8	48	100.0	246	6 ABR54582	Prostate
9	48	100.0	246	7 ADB14478	Human P50
10	48	100.0	252	4 AAE01423	Human sec
11	48	100.0	255	2 AAW85068	Protein e
12	48	100.0	255	3 AAB23268	Human pro
13	48	100.0	255	4 AAU04205	Prostate-
14	48	100.0	255	5 AAO19084	Human pro
15	48	100.0	305	6 ABR54583	Prostate
16	48	100.0	305	7 ADB14479	Human P50
17	48	100.0	355	6 ABR54578	Prostate
18	48	100.0	355	7 ADB14461	Fusion co
19	48	100.0	553	2 AAW69385	Prostate
20	48	100.0	553	3 AAW71869	Amino aci
21	48	100.0	553	3 AAY82002	Human imm
22	48	100.0	553	3 AAB28527	Protein e
23	48	100.0	553	3 ABR54411	Human pro
24	48	100.0	553	4 AAM01117	Human pro
25	48	100.0	553	4 AAU69763	Human pro

26	48	100.0	553	4 AAB74800	Prostate
27	48	100.0	553	4 AAG99002	Human pro
28	48	100.0	553	4 AAG62150	Human P50
29	48	100.0	553	4 ABU71653	Prostate
30	48	100.0	553	4 AAU04961	Human pro
31	48	100.0	553	5 AAU10324	Human pro
32	48	100.0	553	5 ABB95222	Human L1-
33	48	100.0	553	5 AAU82643	Human bre
34	48	100.0	553	5 ABB61900	Prostate
35	48	100.0	553	5 ABR76665	Prostate
36	48	100.0	553	5 ABB77575	Human mas
37	48	100.0	553	6 ABR54334	Prostate
38	48	100.0	553	7 ADB13563	Human pro
39	48	100.0	595	7 AAM01318	Alpha pre
40	48	100.0	710	5 AAW50661	Thioredox
41	48	100.0	1079	4 AAB74830	Prostate
42	48	100.0	1079	4 ABEU71860	Prostate
43	44	91.7	20	4 AAM01145	Human pro
44	44	91.7	20	4 AAU69790	Human pro
45	44	91.7	20	4 AAG99030	Human pro

ALIGNMENTS

RESULT 1

AAM01228
ID AAM01228 standard; protein; 123 AA.

XX AC AAM01228;

XX DT 04-OCT-2001 (first entry)

XX DE P5538 splice variant amino acid P5538-6.

XX KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;

XX KY cytosstatic; gene therapy; metastasis.

XX OS Homo sapiens.

XX PN WO200151633-A2.

XX PD 19-JUL-2001.

XX PF 16-JAN-2001; 2001WO-US001574.

XX PR 14-JAN-2000; 2000US-00483672.

XX XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;

XX WPI; 2001-425873/45.

XX New polynucleotide encoding a prostate-specific protein, for diagnosing,
PT monitoring and treating prostate cancer in a patient and for use in
PT vaccines.

XX Claim 2; Page 463-464; 543pp; English.

XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC the antibodies are also used in the detection of cancer in a patient. The
CC cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC and (II) can be used in vaccines. The antibodies or (I) can be used for
CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate

CC volume. AAH93357 to AAH93944 and AAH01115 to AAH01318 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 48; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
| | | | |
DB 45 GATCLSHSV 53

RESULT 2
AAU69873
ID AAU69873 standard; protein; 123 AA.
XX
XX AAU69873;
AC AAU69873;
XX
DT 30-JAN-2002 (first entry)
XX
DE Human prostate cDNA encoded protein #67.
XX
KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS Homo sapiens.
XX
XX WO200173032-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US009919.
XX
XX 27-MAR-2000; 2000US-00536957.
XX
XX 09-MAY-2000; 2000US-00568100.
XX
XX 12-MAY-2000; 2000US-00570737.
XX
XX 13-JUN-2000; 2000US-00593793.
XX
XX 27-JUN-2000; 2000US-00605783.
XX
XX 09-AUG-2000; 2000US-00636215.
XX
XX 29-AUG-2000; 2000US-00651236.
XX
XX 06-SEP-2000; 2000US-00657279.
XX
XX 02-OCT-2000; 2000US-00679426.
XX
XX 10-OCT-2000; 2000US-00685166.
XX
XX 09-NOV-2000; 2000US-00709729.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
XX WPI: 2001-639232/73.
XX
XX N-PSDB; AAS64041.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer.
XX
XX Claim 2; Page 486-487; 579pp; English.
XX
XX The invention relates to isolated prostate-specific polynucleotides,
CC polypeptides, fusion proteins of the polypeptides, antibodies raised
CC against the polypeptides (or antigenic epitopes derived from them) and
CC antigen-presenting cells expressing the polypeptides. The antibodies are
CC useful for detecting the presence of cancer, especially prostate cancer.
CC The polypeptides, polynucleotides and the antigen-presenting cells are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein, and for inhibiting the development of cancer especially prostate
CC cancer. Compositions comprising the polynucleotide and/or polypeptide are
CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polypeptide of the invention

XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 48; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
| | | | |
DB 45 GATCLSHSV 53

RESULT 3
ABU71764
ID ABU71764 standard; protein; 123 AA.
XX
XX ABU71764;
AC ABU71764;
XX
DT 10-JUN-2003 (first entry)
XX
DE Prostate cancer associated protein #33.
XX
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW immunogen; cancer; prostate specific antigen; PSA;
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW PSMA.
XX
XX Homo sapiens.
OS
XX
XX US2002192763-A1.
XX
XX 19-DEC-2002.
XX
XX 29-JUN-2001; 2001US-00895793.
XX
XX 04-OCT-1999; 99US-0157455P.
XX
XX 04-OCT-2000; 2000US-00679272.
XX
XX 28-MAR-2001; 2001US-00822827.
XX
XX (XUJJ) XU J.
XX
XX (DILL) DILLON D C.
XX
XX (MITC) MITCHAM J L.
XX
XX (HARL) HARLOCKER S L.
XX
XX (JIAN) JIANG Y.
XX
XX (KALO) KALOS M D.
XX
XX (FANG) FANGER G R.
XX
XX (RETT) RETTER M W.
XX
XX (STOL) STOLK J A.
XX
XX (DAYC) DAY C H.
XX
XX (VEDV) VEDVICK T S.
XX
XX (CART) CARTER D.
XX
XX (LISX) LI S X.
XX
XX (WANG) WANG A.
XX
XX (SKEI) SKEIKY Y A W.
XX
XX (HEPL) HEPLER W T.
XX
XX (HEND) HENDERSON R A.
XX
XX (HURA) HURAL J.
XX
XX (MCNE) MCNEILL P D.
XX
XX (HOUG) HOUGHTON R L.
XX
XX (DRAS) Y DE BASSOLS C V.
XX
XX (FOYT) FOY T M.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
PI Mcneill PD, Houghton RL, Y De BassolsCV, Foy TM;
XX
XX WPI: 2001-245062/25.
XX
XX Prostate specific protein and its encoding polynucleotide, useful for the
PT treatment and diagnosis of prostate cancer.
XX
XX Example 13; SEQ ID NO 706; 85pp; English.
PS

XX The invention describes a fusion protein comprising at least one amino
 CC acid sequence of immunogenic portions of any of the 3 sequences not
 CC defined in the specification, or sequences having at least 70 or 90 %
 CC sequence identity to any one of the 35 sequences defined in the USPTO web
 CC site, which is encoded by any of the 4 nucleotide sequences not defined
 CC in the specification. The fusion protein, composition and methods are
 CC useful for diagnosing, preventing and/or treating cancer, particularly
 CC prostate cancer. The proteins are useful as markers to indicate the
 CC presence or absence of cancer. This is the amino acid sequence of a
 CC prostate cancer therapy associated protein. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?docID=US20020192763
 XX

SQ Sequence 123 AA;

Query Match 100.0%; Score 48; DB 4; Length 123;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
 |||||
 DB 45 GATCLSHSV 53

RESULT 4
 ABB95333
 ID ABB95333 standard; protein; 123 AA.

AC ABB95333;

XX 19-JUL-2002 (first entry)

DE Human P553S splice variant encoded protein SEQ ID NO 706.

XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
 KW gene therapy.

XX Homo sapiens.

XX US2002022248-A1.

XX 21-FEB-2002.

XX 12-JAN-2001; 2001US-00759143.

XX 25-FEB-1997; 97US-00806099.

XX 01-AUG-1997; 97US-00904804.

XX 10-FEB-1998; 98US-00020956.

XX 25-FEB-1998; 98US-00030607.

XX 14-JUL-1998; 98US-00115453.

XX 23-SEP-1998; 98US-00159812.

XX 15-JAN-1999; 99US-00232149.

XX 09-APR-1999; 99US-00288946.

XX 13-JUL-1999; 99US-00352616.

XX 12-NOV-1999; 99US-00439313.

XX 18-NOV-1999; 99US-00443286.

XX 14-JAN-2000; 2000US-00483672.

XX 27-MAR-2000; 2000US-00536857.

XX 09-MAY-2000; 2000US-00568100.

XX 12-MAY-2000; 2000US-00570737.

XX 13-JUN-2000; 2000US-00593793.

XX 27-JUN-2000; 2000US-00605783.

XX 10-AUG-2000; 2000US-00636215.

PA (HARL/) HARLOCKER S L.
 PA (JIAN/) JIANG Y.
 PA (KALO/) KALOS M D.
 PA (FANG/) FANGER G R.
 PA (REIT/) RETTER M W.
 PA (STOL/) STOLK J A.
 PA (DAY/) DAY C H.
 PA (VEDV/) VEDVICK T S.
 PA (CART/) CARTER D.
 PA (LISK/) LI S X.
 PA (WANG/) WANG A.
 PA (SKEI/) SKEIKY Y A W.
 PA (HEPL/) HEPLER W T.
 PA (HEND/) HENDERSON R A.
 XX

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2002-255649/30.

DR New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer.

XX Claim 2; SEQ ID NO 706; 87pp; English.

XX The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancers, particularly prostate cancer. The present sequence is a
 CC protein described in the invention

XX Sequence 123 AA;

Query Match 100.0%; Score 48; DB 5; Length 123;

Best Local Similarity 100.0%; Pred. No. 0.58;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCLSHSV 9
 |||||
 DB 45 GATCLSHSV 53

RESULT 5

ABR54445

XX ABR54445 standard; protein; 123 AA.

XX ABR54445;

XX 28-AUG-2003 (first entry)

XX Prostate tumour specific protein sequence SEQ ID 706.

XX Cytostatic; gene therapy; prostate-specific protein; PSP; human;

XX Immune response; prostate cancer.

XX Homo sapiens.

XX WO200289747-A2.

XX 14-NOV-2002.

XX 09-MAY-2002; 2002WO-US014753.

XX 09-MAY-2001; 2001US-00852911.

XX 29-JUN-2001; 2001US-00895814.

XX 10-DEC-2001; 2001US-00012896.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;

PI Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS;

PI Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hural J;

PI McNeill PD, Houghton RL, Vinals Y De BassolsC, Foy TM, Watanabe Y;
PI Deng T;
XX WPI; 2003-167130/16.
DR
XX
XX New prostate-specific proteins and genes, useful in gene therapy,
PT particularly for stimulating an immune response in a patient, or treating
PT prostate cancer in a patient, as well as for diagnosing prostate cancer
PT in a patient.
XX
XX
XX Example 13; Page 555; 691pp; English.
XX
XX The present invention relates to novel prostate-specific proteins (PSP)
CC and their coding sequences. The PSPs and their coding sequences are
CC useful for stimulating an immune response in a patient, or for treating
CC prostate cancer in a patient and for determining, detecting or diagnosing
CC the presence of a cancer in a patient. The present sequence was used to
CC illustrate the invention
XX
SQ Sequence 123 AA;
Query Match 100.0%; Score 48; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCLSHSV 9
Db 45 GATCLSHSV 53
RESULT 6
ADBI4156
ID ADB14156 standard; protein; 123 AA.
AC ADB14156;
XX
XX 18-DEC-2003 (first entry)
DT
DE Human prostate specific cDNA P553S splice variant #4 protein.
KW Human; prostate specific cDNA; cytostatic; immunostimulant; gene therapy;
KW cell therapy; vaccine; T-cell epitope;
KW class I major histocompatibility complex allele; MHC; prostate cancer;
KW tumour; antigen presenting cell.
XX
OS Homo sapiens.
PN US2003185830-A1.
XX
PD 02-OCT-2003.
XX
XX 12-NOV-2002; 2002US-00294025.
XX
XX 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 09-FEB-1998; 98US-00202956.
PR 23-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00152616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443686.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.

PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
PR 12-JAN-2001; 2001US-00759143.
PR 09-FEB-2001; 2001US-00780669.
PR 09-MAY-2001; 2001US-00852911.
PR 29-JUN-2001; 2001US-00895814.
PR 10-DEC-2001; 2001US-00012896.
PR 09-MAY-2002; 2002US-00144678.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Stolk JA, Kalos MD;
XX WPI; 2003-756193/71.
DR N-PSDB; ADB14155.
XX
XX New isolated polypeptide for use in a vaccine for stimulating an immune
PT response, or for treating or diagnosis cancer, preferably prostate
PT cancer.
XX
XX Example 13; Page; 101pp; English.
PS
XX The invention relates to an isolated polypeptide comprising no more than
CC 11-542 amino acids of ADB13563 comprising a sequence ADB14487. The
CC peptides comprise a fragment ADB13563 of that contain naturally processed
CC T-cell epitopes for 3 class I major histocompatibility complex (MHC)
CC alleles. ADB13563 is a polypeptide encoded by a human prostate specific
CC cDNA, one of 648 disclosed as new. Also included are nucleic acids
CC encoding the proteins and peptides, expression vectors, a host cell
CC transformed with the vector, an isolated antibody (or antigen binding
CC fragment) that specifically binds to the protein or peptide, detecting
CC the presence of a cancer in a patient (comprising contacting a patient
CC sample with a binding agent that binds to the peptides or a polypeptide
CC appearing as ADB13563, detecting the amount of polypeptide that binds to
CC the agent and comparing the amount of polypeptide to a predetermined cut-
CC off value to determine the presence of cancer), a fusion protein
CC comprising the peptides or proteins, stimulating or expanding T cells
CC specific for a tumour protein comprising contacting T cells with the
CC peptides or the isolated T cell population, treating prostate cancer in a
CC patient comprising administering a composition comprising the peptides,
CC nucleic acids, antibodies or compounds, determining the presence of a
CC cancer in a patient and treating prostate cancer in a patient comprising
CC incubating cluster of differentiation (CD4+) and/or CD8+ T cells isolated
CC from a patient with the peptides or antigen presenting cells that express
CC (the peptides so that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The peptides (or an oligonucleotide
CC that hybridises to nucleic acid encoding them), is used to detect the
CC presence of cancer in a patient. The peptides, nucleic acids encoding, or
CC antigen-presenting cells expressing the nucleic acid, are used to
CC stimulate or expand T cells specific for a tumour protein. The peptides,
CC nucleic acids, antibodies, fusion proteins, T cell populations or antigen
CC presenting cells are used to stimulate an immune response or treat
CC prostate cancer in a patient. The present sequence is a prostate specific
CC protein of the invention. Note: Except where otherwise indicated, the
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030185830.
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 48; DB 7; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCLSHSV 9
Db 45 GATCLSHSV 53

RESULT 7
AAY54369
ID AAY54369 standard; protein; 231 AA.

XX AC AAY54369;
XX DT 06-APR-2000 (first entry)
XX AC
XX DE Amino acid sequence of a prostate cancer-associated protein.
XX DE Prostate cancer-associated gene; Incyte clone 1864683; bone cancer;
XX KW cell proliferation; cancer; adrenal gland cancer; bladder cancer;
XX KW prostate cancer.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT Peptide 1..47
XX FT /note= "signal peptide"
XX FT Modified-site 100
XX FT /note= "potential casein kinase II phosphorylation site"
XX FT Modified-site 142
XX FT /note= "potential casein kinase II phosphorylation site"
XX FT Modified-site 147
XX FT /note= "potential protein kinase C phosphorylation site"
XX FT Domain 188..209
XX FT /note= "potential transmembrane domain"
XX FN WO9967384-A2.
XX PD 29-DEC-1999.
XX PP 15-JUN-1999; 99WO-US013524.
XX PP 22-JUN-1998; 98US-00102615.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Walker MG, Volkmuth W, Klingler TM, Sprinzak EA;
XX DR WPI; 2000-126631/11.
XX DR N-PSDB; AAZ45677.
XX PT Identifying biomolecules for the diagnosis and treatment of diseases
XX PT associated with cell-proliferation.
XX PS Claim 10; Page 50-51; 52pp; English.
XX CC The present sequence is encoded by a prostate cancer-associated gene,
XX CC whose cDNA is represented by Incyte clone number 1864683. The cDNA
XX CC sequence is used in the method of the invention. The specification
XX CC describes a method for identifying biomolecules for the diagnosis or
XX CC treatment of diseases associated with cell proliferation. The method
XX CC comprises examining polynucleotides, consisting of prostate cancer-
XX CC specific genes, and genes of unknown function, expressed in cDNA
XX CC libraries. The patterns of both gene sets are compared to identify genes
XX CC of unknown function with similar expression patterns to the prostate
XX CC cancer-specific genes. The biomolecules identified by the method form
XX CC pharmaceutical compositions useful for the diagnosis and treatment of
XX CC diseases associated with cell proliferation. Such diseases include cancer
XX CC of the adrenal gland, bladder and bone, but especially prostate cancer.
XX CC The method may also be applied using other disease-specific genes. The
XX CC prostate cancer-specific genes facilitate the diagnosis and treatment of
XX CC cell proliferation disorders
XX SQ Sequence 231 AA;
Query Match 100.0%; Score 48; DB 3; Length 231;
Best Local Similarity 100.0%; Pred. NO. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCLSHSV 9
DB 45 GATCLSHSV 53

RESULT 8
ABR54582
ID ABR54582 standard; protein; 246 AA.
XX AC
XX AC ABR54582;
XX DT 28-AUG-2003 (first entry)
XX DE Prostate tumour-related protein SEQ ID 1028.
XX KW Cytostatic; gene therapy; prostate-specific protein; PSP; human;
XX KW immune response; prostate cancer.
XX OS Homo sapiens.
XX PN WO200289747-A2.
XX PD 14-NOV-2002.
XX PP 09-MAY-2002; 2002WO-US014753.
XX PR 09-MAY-2001; 2001US-00852911.
XX PR 29-JUN-2001; 2001US-00895814.
XX PR 10-DEC-2001; 2001US-00012896.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;
XX PI Kalos MD, Fanger GR, Retter MW, Scolk JA, Day CH, Vedvick TS;
XX PI Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hural J;
XX PI McNeill PD, Houghton RL, Vinals Y De Bassolac, Foy TM, Watanabe Y;
XX PI Deng T;
XX DR WPI; 2003-167130/16.
XX PT New prostate-specific proteins and genes, useful in gene therapy,
XX PT particularly for stimulating an immune response in a patient, or treating
XX PT prostate cancer in a patient, as well as for diagnosing prostate cancer
XX PT in a patient.
XX PS Claim 2; Page 688-689; 691pp; English.
XX CC The present invention relates to novel prostate-specific proteins (PSP)
XX CC and their coding sequences. The PSPs and their coding sequences are
XX CC useful for stimulating an immune response in a patient, or for treating
XX CC prostate cancer in a patient and for determining, detecting or diagnosing
XX CC the presence of a cancer in a patient. The present sequence was used to
XX CC illustrate the invention
XX SQ Sequence 246 AA;
Query Match 100.0%; Score 48; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. NO. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCLSHSV 9
DB 60 GATCLSHSV 68
RESULT 9
ABR14478
ID ABR14478 standard; protein; 246 AA.
XX AC
XX AC ABR14478;
XX DT 18-DEC-2003 (first entry)
XX DE Human P501S deletion mutant, P501S D.
XX KW Human; prostate specific cDNA; cytostatic; immunostimulant; gene therapy;
XX KW cell therapy; vaccine; T-cell epitope;
XX KW class I major histocompatibility complex allele; MHC; prostate cancer;

PA (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;
PI WPI; 2001-308779/32.
XX
XX New nucleic acid encoding one of 21 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions, such
PT as autoimmune disease and cancer, and used as a food additive or
PT preservative.
XX
XX Disclosure; Page 476-477; 490pp; English.
XX
XX AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC protein genes, and AAD01352-AAE01413 represent the proteins they encode.
CC AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 21 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies, disease),
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein variant referred to
CC in the disclosure of the invention
XX
XX Sequence 252 AA;
SQ
Query Match 100.0%; Score 48; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATCLSHSV 9
Db 45 GATCLSHSV 53
|||||
RESULT 11
AAW85068
ID AAW85068 standard; protein; 255 AA.
XX
XX AAW85068;
AC
XX 12-FEB-1999 (first entry)
DE
XX Protein encoded by the consensus sequence of the PS108 gene.
XX
XX PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;
XX prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;
XX drug screening; Gene therapy.
XX
XX Homo sapiens.
OS
XX WO9850567-A1.
XX
XX 12-NOV-1998.
PD
Query Match 100.0%; Score 48; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATCLSHSV 9
Db 69 GATCLSHSV 77
|||||
RESULT 12
AAB29268
ID AAB29268 standard; protein; 255 AA.
XX
XX AAB29268;
AC
XX 07-FEB-2001 (first entry)
DE
XX Human prostate-related PS108 partial protein sequence SEQ ID NO: 36.
XX
XX Human; prostate cancer; PS108; antibody; tumour; metastasis.
XX
XX Homo sapiens.
OS
XX US6130043-A.
XX
XX 10-OCT-2000.
PD
XX
XX 01-MAY-1998; 98US-00071710.
XX
XX 02-MAY-1997; 97US-00850713.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Cohen M, Colpitts TL, Friedman PN, Gordon J, Granados EN;
XX Billing-Medel PA, Klass MR, Kratochvil JD, Stroupe SD, Yu H;
XX Kratochvil JD, Russell JC, Hodges SC;
XX

XX
PF 01-MAY-1998; 98WO-US008930.
XX
XX 02-MAY-1997; 97US-00850713.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
XX Russell JC, Stroupe SD;
XX
XX WPI: 1999-034731/03.
XX
XX N-PSDB; AAV71181.
XX
XX New isolated prostate-specific polynucleotides - used to develop products
PT for the diagnosis and treatment of prostate diseases, e.g. benign
PT hyperplasia, prostatic or prostate cancer.
XX
XX Claim 17; Page 99-100; 122pp; English.
XX
XX The present sequence is encoded by the consensus sequence for a PS108
CC gene. The sequence is derived from overlapping clones AAV71186-79. The
CC clone sequences are PS108 gene-specific. They are used in the method of
CC the invention. The specification describes a method for detecting the
CC presence of a target PS108 polynucleotide in a test sample. The method
CC comprises contacting the test sample with at least 1 PS108-specific
CC polynucleotide or complement, and detecting the presence of the target
CC PS108 polynucleotide. The products can be used for detecting, diagnosing,
CC staging, monitoring, prognosticating, in vivo imaging, preventing or
CC treating, or determining predisposition to diseases or conditions of the
CC prostate such as benign prostatic hyperplasia (BPH), prostatitis,
CC prostatic intraepithelial neoplasia (PIN) and cancer. In particular the
CC products can be used in drug screening and gene therapy
XX
XX Sequence 255 AA;
SQ
Query Match 100.0%; Score 48; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATCLSHSV 9
Db 69 GATCLSHSV 77
|||||

DR WPI; 2000-655655/63.
 XX Methods for detecting target prostate-specific polynucleotides or
 PT diseases of the prostate (e.g. prostate cancer), comprising detecting the
 PT presence of any of PS108 nucleic acid sequences in a test sample.
 XX Example 1; Col 85-88; 55pp; English.
 XX The present invention is related to a number of partial coding and
 CC protein sequences for the human prostate tissue protein PS108. These
 CC sequences can be used in the diagnosis and prognosis of prostate
 CC diseases, particularly prostate cancer. They can also be used to produce
 CC antibodies which can be used in treatment. The present sequence is one of
 CC the PS108 partial protein sequences
 XX
 SQ Sequence 255 AA;
 Query Match 100.0%; Score 48; DB 3; Length 255;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCLSHSV 9
 |||||
 Db 69 GATCLSHSV 77
 RESULT 13
 AAU04205
 ID AAU04205 standard; protein; 255 AA.
 AC AAU04205;
 DT 23-OCT-2001 (first entry)
 DE Prostate-specific PS108 polypeptide #1.
 XX Prostate; PS108; immunogen; drug screening; image localisation;
 KW diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;
 KW expressed sequence tag; EST.
 XX Homo sapiens.
 OS
 XX US6252047-B1.
 XX 26-JUN-2001.
 XX 15-MAR-2000; 2000US-00525397.
 XX 02-MAY-1997; 97US-00850713.
 XX 01-MAY-1998; 98US-00071710.
 XX (ABBO) ABBOTT LAB.
 XX Billigell PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD, Yu H;
 XX WPI; 2001-424488/45.
 XX N-PSDB; AAS07155.
 XX Novel PS108 polypeptide useful in assays for detecting antibodies to
 PT prostate tissue, and as immunogens to produce PS108 antibodies.
 XX
 PS Claim 1; Col 85-87; 55pp; English.
 XX The sequence represents the amino acid sequence of prostate-specific
 CC PS108 #1, encoded by PS108 expressed sequence tag (EST) consensus
 CC sequence. The sequence was produced from overlapping PS108 ESTs sequences
 CC to produce a full length consensus sequence. This sequence was then used
 CC to produce the PS108 polypeptide which is useful in assays for detecting
 CC antibodies to prostate tissue, and as an immunogen to produce antibodies.
 CC The polypeptide is useful for screening compounds which specifically bind
 CC to the polypeptide and for screening for drugs, compounds, or any other

CC agent which can be used to treat diseases associated with PS108. The
 CC antibody is useful to detect, or for image localisation of, PS108 antigen
 CC in a patient, for detecting or diagnosing a disease or condition, as
 CC delivery agents for therapeutic agents as well as for diagnostic tests
 CC and for screening for diseases or conditions associated with PS108, and
 CC especially cancer. The antibody is also useful for generating chimeric
 CC antibodies for therapeutic use, for inhibiting the biological activity of
 CC PS108, in therapy (for e.g. to treat prostate tissue disease including
 CC polypeptide in a test sample which shares one or more antigenic
 CC determinants with the PS108 polypeptide
 XX
 SQ Sequence 255 AA;
 Query Match 100.0%; Score 48; DB 4; Length 255;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCLSHSV 9
 |||||
 Db 69 GATCLSHSV 77
 RESULT 14
 AA019084
 ID AA019084 standard; protein; 255 AA.
 AC AA019084;
 XX 22-NOV-2002 (first entry)
 DE Human prostate-specific PS108 protein SEQ ID NO: 36.
 XX Human; PS108; prostate cancer; prostate specific sequence; prostate;
 KW cytosstatic; gene therapy.
 XX Homo sapiens.
 OS
 XX US2002086301-A1.
 XX 04-JUL-2002.
 XX 25-APR-2001; 2001US-00841894.
 XX 02-MAY-1997; 97US-00850713.
 XX 01-MAY-1998; 98US-00071710.
 XX (BILL) BILLINGEL P A.
 PA (COHE) COHEN M.
 PA (COLP) COLPITTS T L.
 PA (FRIE) FRIEDMAN P N.
 PA (GORD) GORDON J.
 PA (GRAN) GRANADOS E N.
 PA (HODG) HODGES S C.
 PA (KLAS) KLASS M R.
 PA (KRAT) KRATOCHVIL J D.
 PA (ROBE) ROBERTS-RAPP L A.
 PA (RUSSE) RUSSELL J C.
 PA (STRO) STROUPE S D.
 PA (YUHH) YU H.
 XX Billigell PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp LA;
 PI Russell JC, Stroupe SD, Yu H;
 XX WPI; 2002-665428/71.
 XX Novel PS108 polynucleotides derived from the PS108 gene, useful for
 PT detecting, diagnosing, staging, monitoring, prognosticating, in vivo
 PT imaging, preventing or treating diseases and conditions of prostate e.g.,
 XX prostate cancer.
 XX Claim 17; Page 46; 59pp; English.

XX The present invention relates to polynucleotide sequences derived from
CC the human PS108 gene and capable of selectively hybridising to the PS108
CC gene. The sequences can be used to detect the presence of a target PS108
CC polynucleotide in a test sample, and to detect, monitor, prevent or treat
CC diseases and conditions of the prostate, such as prostate cancer. The
CC present sequence is a PS108 protein fragment
XX
SQ Sequence 255 AA;

Query Match 100.0%; Score 48; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCLSHSV 9
Db 69 GATCLSHSV 77

RESULT 15
ABR54583
ID ABR54583 standard; protein; 305 AA.

XX AC ABR54583;

XX 28-AUG-2003 (first entry)

XX Prostate tumour-related protein SEQ ID 1029.

XX Cytostatic; gene therapy; prostate-specific protein; PSP; human;
KW immune response; prostate cancer.

XX Homo sapiens.

XX WO200289747-A2.

XX 14-NOV-2002;

XX 09-MAY-2002; 2002WO-US014753.

XX 09-MAY-2001; 2001US-00852911.

XX 29-JUN-2001; 2001US-00895814.

XX 10-DEC-2001; 2001US-00012896.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;
PI Kalos MD, Fanger GR, Retter MW, Stoik JA, Day CH, Vedvick TS;
PI Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Rural J;
PI McNeill PD, Houghton RL, Vinals Y De BassolsC, Foy TM, Watanabe Y;
PI Deng T;

XX WPI; 2003-167130/16.

XX New prostate-specific proteins and genes, useful in gene therapy,
PT particularly for stimulating an immune response in a patient, or treating
PT prostate cancer in a patient, as well as for diagnosing prostate cancer
PT in a patient.

XX Claim 2; Page 689-690; 691pp; English.

XX The present invention relates to novel prostate-specific proteins (PSP)
CC and their coding sequences. The PSPs and their coding sequences are
CC useful for stimulating an immune response in a patient, or for treating
CC prostate cancer in a patient and for determining, detecting or diagnosing
CC the presence of a cancer in a patient. The present sequence was used to
CC illustrate the invention

XX Sequence 305 AA;

Query Match 100.0%; Score 48; DB 6; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCLSHSV 9
Db 119 GATCLSHSV 127

Search completed: August 25, 2004, 02:27:08
Job time : 2.78648 secs

This Page Blank (uspto)


```

RESULT 2
US-09-759-143-706
; Sequence 706, Application US/09759143
; Patent No. US200202248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-706

Query Match          100.0%; Score 48; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 3
US-09-780-669-706
; Sequence 706, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669

Query Match          100.0%; Score 48; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 4
US-09-822-827-706
; Sequence 706, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-706

Query Match          100.0%; Score 48; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 5
US-09-895-793-706
; Sequence 706, Application US/09895793
; Patent No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.

```



```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-706

Query Match      100.0%; Score 48; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 6
US-09-895-814-706
; Sequence 706, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Wedvick, Thomas S.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-706

Query Match      100.0%; Score 48; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 7
US-10-012-896-706
```

```
; Sequence 706, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Wedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-706

Query Match      100.0%; Score 48; DB 13; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GATCLSHSV 9
Db      45 GATCLSHSV 53

RESULT 8
US-10-144-678A-706
; Sequence 706, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Wedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
```

```

; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-706

```

```

Query Match      100.0%; Score 48; DB 14; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

```

```

Qy 1 GATCLSHSV 9
   |||||
Db 45 GATCLSHSV 53

```

```

RESULT 9
US-10-294-025-706
; Sequence 706, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-706

```

```

Query Match      100.0%; Score 48; DB 14; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.6; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

```

```

Qy 1 GATCLSHSV 9
   |||||
Db 45 GATCLSHSV 53

```

```

RESULT 10
US-10-144-678A-1028
; Sequence 1028, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.

```

```

; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1028
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-1028

```

```

Query Match      100.0%; Score 48; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

```

```

Qy 1 GATCLSHSV 9
   |||||
Db 50 GATCLSHSV 68

```

```

RESULT 11
US-10-294-025-1028
; Sequence 1028, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1028
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-1028

```

```

Query Match      100.0%; Score 48; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

```

```

Qy 1 GATCLSHSV 9
   |||||
Db 60 GATCLSHSV 68

```

```

RESULT 12
US-09-841-894-36
; Sequence 36, Application US/09841894
; Publication No. US20020086301A1
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.

```

FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KEATCHEVIL, JON D.
ROBERTS-RAPF, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,894
FILING DATE: 25-Apr-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/071,710
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:

LENGTH: 255 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 36:

US-09-841-894-36
Query Match 100.0%; Score 48; DB 12; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
DB 69 GATCLSHSV 77

RESULT 13
US-10-144-678A-1029

Sequence 1029, Application US/10144678A
Publication No. US20030157089A1

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugui

APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.

APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Desg. Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144,678A

CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1029

LENGTH: 305
TYPE: PRT

ORGANISM: Homo sapiens
US-10-144-678A-1029

Query Match 100.0%; Score 48; DB 14; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
DB 119 GATCLSHSV 127

RESULT 14
US-10-294-025-1029

Sequence 1023, Application US/10294025
Publication No. US20030185830A1

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Stolck, John A.
APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C29
CURRENT APPLICATION NUMBER: US/10/294,025

CURRENT FILING DATE: 2002-11-12
NUMBER OF SEQ ID NOS: 1038

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1029

LENGTH: 305
TYPE: PRT

ORGANISM: Homo sapiens
US-10-294-025-1029

Query Match 100.0%; Score 48; DB 14; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
DB 119 GATCLSHSV 127

RESULT 15
US-10-012-896-1011

Sequence 1011, Application US/10012896
Publication No. US20020183251A1

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

```

; APPLICANT: Jiang, Yudi
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 1011
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-1011

```

```

Query Match      100.0%; Score 48; DB 13; Length 355;
Best Local Similarity 100.0%; Pred. NO. 1.7;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GATCLSHSV 9
DB      185 GATCLSHSV 193

```

```

Search completed: August 25, 2004, 02:34:07
Job time : 2.64235 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:00:53 ; Search time 0.800712 Seconds
(without alignments)
1081.193 Million cell updates/sec

Title: US-09-593-793A-113_COPY_367_375

Perfect score: 48

Sequence: 1 GATCLSHSV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	77.1	310	2 A55768	asparaginyl-peptid
2	37	77.1	669	2 T32512	hypothetical prote
3	36	75.0	308	2 T37125	epidermal growth f
4	36	75.0	1208	2 T34469	hypothetical prote
5	35	72.9	511	1 S31308	aldohyde dehydroge
6	34	70.8	156	2 S54619	hypothetical prote
7	34	70.8	311	2 F69820	conserved hypotet
8	34	70.8	898	2 A40114	fasciclin II precu
9	34	70.8	1545	2 T26599	hypothetical prote
10	33	68.8	170	2 AF2621	polypeptide deform
11	33	68.8	170	2 F97403	polypeptide deform
12	33	68.8	187	2 AF3542	formylmethionine d
13	33	68.8	302	2 S59853	DNA-binding protei
14	33	68.8	357	2 T41716	conserved hypotet
15	33	68.8	383	2 S53716	delta-like homeoti
16	33	68.8	385	2 A54785	preadipocyte facto
17	33	68.8	385	2 S53718	homeotic protein d
18	33	68.8	443	1 Q85YET	hypothetical prote
19	33	68.8	627	2 A83408	phenazine biosynth
20	33	68.8	1001	2 S30385	G9a protein - huma
21	33	68.8	1076	2 D82083	carbamoyl-phosphat
22	33	68.8	1216	2 T26104	hypothetical prote
23	33	68.8	1227	2 T49953	hypothetical prote
24	32	66.7	64	2 E69172	sensory transducti
25	32	66.7	148	2 D81852	conserved hypotet
26	32	66.7	174	2 A81888	hypothetical prote
27	32	66.7	246	2 JQ1472	trypsin (EC 3.4.21
28	32	66.7	246	2 JQ1471	trypsin (EC 3.4.21
29	32	66.7	262	1 S17705	indole-3-glycerol-

30 32 66.7 270 2 T46856 indole-3-glycerol-
31 32 66.7 277 1 D35115 indole-3-glycerol-
32 32 66.7 294 2 A72362 cell division prot
33 32 66.7 468 2 T49682 hypothetical prote
34 32 66.7 638 2 AC1123 B. subtilis fold p
35 32 66.7 638 2 A81483 B. subtilis fold p
36 32 66.7 705 2 A35621 spore germination
37 32 66.7 775 2 T37837 probable signal tr
38 32 66.7 1106 2 T18739 hypothetical prote
39 32 66.7 1672 2 T46237 hypothetical prote
40 32 66.7 2023 2 T13154 polycomb protein e
41 32 66.7 2731 1 VFIHQH genome polyprotein
42 32 66.7 2733 2 S15760 genome polyprotein
43 31 64.6 99 2 F82824 hypothetical prote
44 31 64.6 163 2 H83168 conserved hypotet
45 31 64.6 166 2 AB0371 conserved hypotet

ALIGNMENTS

RESULT 1

A55768
asparaginyl-peptide amidohydrolase (EC 3.5.1.-) - pig
N/Alternate names: asparagine amidohydrolase
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 05-Nov-1999
C/Accession: A55768
R/Stewart, A.E.; Arfin, S.M.; Bradshaw, R.A.

J. Biol. Chem. 270, 25-28, 1995
A/Title: The sequence of porcine protein NH-2-terminal asparagine amidohydrolase. A new
A/Reference number: A55768; MUID:95113832; PMID:7814382

A/Accession: A55768

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-310 <STR>

A/Cross-references: GB:U17062; NID:G955949; PIDN:AAA65019.1; PID:G595950

C/Keywords: hydrolase

F/2-310/Product: asparaginyl-peptide amidohydrolase #status predicted <MAT>

Query Match 77.1%; Score 37; DB 2; Length 310;

Best Local Similarity 85.7%; Pred No. 14;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSH 7

Db 86 GATCLTH 92

RESULT 2

T32512
hypothetical protein C44B12.4 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T32512

R/Fin-Wollam, A.

submitted to the EMBL Data Library, December 1997

A/Description: The sequence of C. elegans cosmid C44B12.

A/Reference number: Z21183

A/Accession: T32512

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-669 <FIN>

A/Cross-references: EMBL:AF036692; PIDN:AAB88327.1; GSPDB:GN00022; CESP:C44B12.4

A/Experimental source: strain Bristol N2; clone C44B12

C/Genetics:

A/Gene: CESP:C44B12.4

A/Map position: 4

A/Introns: 35/3; 89/2; 133/3; 164/1; 213/3; 288/2; 327/2; 361/3; 380/1; 448/2; 638/2

Query Match 77.1%; Score 37; DB 2; Length 669;

Best Local Similarity 66.7%; Pred. No. 27;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9
 DB 62 GTCLSHSI 70
 RESULT 3
 JC7125
 epidermal growth factor-like homeotic protein PREP-1 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 02-Aug-2002
 C:Accession: JC7125
 R:Fahrenkrug, S.C.; Freking, B.A.; Smith, T.P.L.
 Biochem. Biophys. Res. Commun. 264, 662-667, 1999
 A:Title: Genomic organization and genetic mapping of the bovine PREP-1 gene.
 A:Reference number: JC7125; MUID:20012729; PMID:10543989
 A:Accession: JC7125
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-308 <FAH>
 A:Cross-references: GB:AF181831; NID:g5932411
 C:Genetics: Dlk
 A:Gene: Dlk
 A:Map position: 21 (77.5 cM)
 C:Superfamily: preadipocyte factor 1; EGF homology
 Query Match 75.0%; Score 36; DB 2; Length 308;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GATCLSHS 8
 DB 220 GGICLQHS 227
 RESULT 4
 T34469
 hypothetical protein W03A3.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34469
 R:Miller, N.
 submitted to the EMBL Data Library, February 1996
 A:Description: The sequence of C. elegans cosmid W03A3.
 A:Reference number: Z21531
 A:Accession: T34469
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1208 <MIL>
 A:Cross-references: EMBL:U50184; PIDN:AA93325.1; GSPDB:GN00021; CESP:W03A3.2
 A:Experimental source: strain Bristol N2; clone W03A3
 C:Genetics:
 A:Gene: CESP:W03A3.2
 A:Map position: 3
 A:Introns: 93/3; 119/3; 197/3; 221/1; 267/2; 302/2; 555/1; 589/3; 666/1; 692/1; 744/3; 9
 Query Match 75.0%; Score 36; DB 2; Length 1208;
 Best Local Similarity 85.7%; Pred. No. 69;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GATCLSH 7
 DB 275 GVTCLSH 281
 RESULT 5
 S31308
 aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 2 precursor, mitochondrial - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 31-Jan-1997 #text_change 03-Jun-2002
 C:Accession: S31308
 R:Thielen, J.
 submitted to the EMBL Data Library, October 1992

A:Reference number: S31308
 A:Accession: S31308
 A:Molecule type: DNA
 A:Residues: 1-511 <THI>
 A:Cross-references: EMBL:Z17314; NID:g3361; PIDN:CAA78962.1; PID:g3362
 C:Comment: This form is probably mitochondrial.
 C:Genetics:
 A:Gene: ALD2
 C:Function:
 A:Description: catalyzes the oxidation of an aldehyde to an acid using NAD- and water
 A:Note: enzymes with this activity are involved in diverse metabolic pathways in various
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 C:Keywords: alcohol metabolism; mitochondrion; NAD; oxidoreductase
 F:1-21/Domain: transit peptide (mitochondrion) #status predicted <TNP>
 F:22-511/Product: aldehyde dehydrogenase (NAD+) 2 #status predicted <MAT>
 F:80-351/Domain: aldehyde dehydrogenase homology <ALDD>
 F:297,331/Active site: Glu, Cys #status predicted
 F:483/Binding site: NAD (Cys) #status predicted
 Query Match 72.9%; Score 35; DB 1; Length 511;
 Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GATCLSH 7
 DB 258 GAACLSH 264
 RESULT 6
 S54619
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
 C:Accession: S54619; S66879
 R:de Haan, M.; Maarse, A.C.; Grivell, L.A.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: S54617
 A:Accession: S54619
 A:Molecule type: DNA
 A:Residues: 1-156 <DBH>
 A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123
 R:de Haan, M.; Grivell, L.A.; Maarse, A.C.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66877
 A:Accession: S66879
 A:Molecule type: DNA
 A:Residues: 1-156 <DEW>
 A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR01
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0005539
 A:Map position: 1SR
 C:Superfamily: hypothetical protein YOR013w
 Query Match 70.8%; Score 34; DB 2; Length 156;
 Best Local Similarity 85.7%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GATCLSH 7
 DB 23 GATCLH 29
 RESULT 7
 Z69820
 conserved hypothetical protein yhbB - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: Z69820
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallez, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krohn, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A. *Bacillus subtilis*.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69820
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-311 <KUN>
A;Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12720.1; PID:G2633215
A;Experimental source: strain 168
A;Genetics:
A;Gene: yhbB

Query Match 70.8%; Score 34; DB 2; Length 311;
Best Local Similarity 55.8%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Gaps 0;

Qy 1 GATCLSHSV 9
| | | | | | | | | |
Db 72 GTTCLSVTI 80

RESULT 8
A40114
C;Species: *Schistocerca americana* (American bird grasshopper)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 17-Mar-2000
A;Accession: A40114; B31817
R;Harrison, A.L.; Goodman, C.S.
Science 242, 700-708, 1988
A;Title: Growth cone guidance in insects: fasciclin II is a member of the immunoglobulin
A;Reference number: A40114; MUID:89043938; PMID:3187519
A;Accession: A40114
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-898 <HAR>
R;Snow, P.N.; Zinn, K.; Harrelson, A.L.; McAllister, L.; Schilling, J.; Bastiani, M.J.;
Proc. Natl. Acad. Sci. U.S.A. 85, 5291-5295, 1988
A;Title: Characterization and cloning of fasciclin I and fasciclin II glycoproteins in b
A;Reference number: A94202; MUID:88276943; PMID:2833942
A;Accession: B31817
A;Molecule type: mRNA
A;Residues: 423-436 <SNO>
A;Cross-references: GB:J03789; NID:G160848; PID:G160849
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C;Keywords: transmembrane protein

Query Match 70.8%; Score 34; DB 2; Length 898;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCLSHSV 9
| | | | | | | | | |
Db 449 TCLAHSI 455

RESULT 9
T26589
hypothetical protein Y32B12B.4 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A;Accession: T26589
R;Gardner, A.
submitted to the EMBL Data Library, September 1998

A;Reference number: 220239
A;Accession: T26589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1545 <NIL>
A;Cross-references: EMBL:AL031632; PIDN:CAA21009.1; GSPDB:GN000023; CESP:Y32B12B.4
A;Experimental source: clone Y32B12B
A;Genetics:
A;Gene: CESP:Y32B12B.4
A;Map position: 5
A;Introns: 49/2; 311/2; 879/2; 1116/2; 1257/3; 1404/2; 1545/2

Query Match 70.8%; Score 34; DB 2; Length 1545;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATCLSHSV 9
| | | | | | | | | |
Db 1494 ATCTHSI 1501

RESULT 10
AF2621
polypeptide deformylase def [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C;Species: *Agrobacterium tumefaciens*
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
A;Accession: AF2621
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tac, Y.; Biddle, P.; Jung, M.; Krepsan, W.; Perry, M.; Gordon-Kamm,
Scher, E.W.
A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <KUR>
A;Cross-references: GB:AB008698; PIDN:AAU41388.1; PID:G17738706; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
A;Genetics:
A;Gene: def
A;Map position: circular chromosome
C;Superfamily: polypeptide deformylase

Query Match 68.8%; Score 33; DB 2; Length 170;
Best Local Similarity 62.5%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ATCLSHSV 9
| | | | | | | | | |
Db 131 ATCLQHEI 138

RESULT 11
F97403
polypeptide deformylase (AF213822) [imported] - *Agrobacterium tumefaciens* (strain C58, (C
C;Species: *Agrobacterium tumefaciens*
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
A;Accession: F97403
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tu*
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <KUR>
A;Cross-references: GB:AB007869; PIDN:AAK86183.1; PID:G15155278; GSPDB:GN00169
A;Genetics:
A;Gene: AGR_C_640

A:Map position: circular chromosome
C:Superfamily: polypeptide deformylase

Query Match 68.8%; Score 33; DB 2; Length 170;
Best Local Similarity 62.5%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCLSHSV 9
|||||
DB 131 ATCLQHEI 138

RESULT 12

AF3542
formylmethionine deformylase (EC 3.5.1.31) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C:Accession: AF3542
R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, L.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3542
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-187 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53505.1; PID:gl7984410; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10264
A:Map position: II
C:Superfamily: polypeptide deformylase
C:Keywords: hydrolase

Query Match 68.8%; Score 33; DB 2; Length 187;
Best Local Similarity 62.5%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCLSHSV 9
|||||
DB 143 ATCLQHEI 150

RESULT 13

S59853
DNA-binding protein Dof3 - maize (fragment)
C:Species: Zea mays (maize)
C>Date: 15-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 29-Oct-1999
C:Accession: S59853; S59851
R:Shuichi, Y.
Submitted to the EMBL Data Library, June 1994
A:Reference number: S59852
A:Accession: S59853
A:Molecule type: mRNA
A:Residues: 1-302 <SHU>
A:Cross-references: EMBL:X79935; NID:gl061307; PIDN:CAA56288.1; PID:gl061308
R:Yanagisawa, S.
Nucleic Acids Res. 23, 3403-3410, 1995
A:Title: A novel DNA-binding domain that may form a single zinc finger motif.
A:Reference number: S59850; MUID:96032831; PMID:7567449
A:Accession: S59851
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 27-99 <YAN>
A:Cross-references: EMBL:X79935
C:Keywords: zinc finger

Query Match 68.8%; Score 33; DB 2; Length 302;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9

DB 176 GAACYPHSV 184

RESULT 14

T41716
conserved hypothetical protein SPAC821.05 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41716
R:Rieger, M.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z22012
A:Accession: T41716
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-357 <RI3>
A:Cross-references: EMBL:AL121770; PIDN:CAB57439.1; GSPDB:GN00066; SPDB:SPAC821.05
A:Experimental source: strain 972h-; cosmid c821
C:Genetics:
A:Gene: SPDB:SPAC821.05
A:Map position: 1
A:Introns: 9/2; 45/1

Query Match 68.8%; Score 33; DB 2; Length 357;
Best Local Similarity 75.0%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCLSHSV 9
|||||
DB 200 ATCLLSL 207

RESULT 15

S53716
delta-like homeotic protein dlk, long splice form precursor - human
N:Alternate names: fetal antigen 1 (FAL)
N:Contains: delta-like homeotic protein dlk, short splice form
C:Species: Homo sapiens (man)
C>Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 02-Aug-2002
C:Accession: S53716; S53717; S71068; B45484; S48713; A44549; S31973; S31974
R:Lee, Y.L.; Helman, L.; Hoffman, T.; Laborda, J.
Biochim. Biophys. Acta 1261, 223-232, 1995
A:Title: dlk, pG2 and Pref-1 mRNAs encode similar proteins belonging to the BGF-like superfamily
A:Reference number: S53716; MUID:95226449; PMID:7711066
A:Accession: S53716
A:Molecule type: mRNA
A:Residues: 1-383 <LEE1>
A:Cross-references: EMBL:U15979; NID:g562105; PIDN:AAA75364.1; PID:g562106
A>Note: the authors translated the codon CAG for residue 46 as His and CCF for residue 47 as Phe
A:Note: the sequence in GenBank entry HSU15979 has a 3 base deletion mutation in the region corresponding to residues 348-350
A:Accession: S53717
A:Molecule type: mRNA
A:Residues: 1-228-302-383 <LEE2>
A:Cross-references: EMBL:U15981; NID:g562109; PIDN:AAA75365.1; PID:g562110
R:Laborda, J.
submitted to the EMBL Data Library, October 1994
A:Reference number: S71068
A:Accession: S71068
A:Molecule type: mRNA
A:Residues: 1-344-346-383 <LAB1>
A:Cross-references: EMBL:U15979; NID:g562105; PIDN:AAA75364.1; PID:g562106
R:Laborda, J.; Sausville, E.A.; Hoffman, T.; Notario, V.
J. Biol. Chem. 268, 3817-3820, 1993
A:Title: dlk, a putative mammalian homeotic gene differentially expressed in small cell lung carcinoma cell lines
A:Reference number: A45484; MUID:93179372; PMID:8095043
A:Accession: B45484
A:Molecule type: mRNA
A:Residues: 1-45 'HV', 48-383 <LAB2>
A:Cross-references: EMBL:Z12172; NID:g38478; PIDN:CAA78163.1; PID:g38479
A>Note: sequence extracted from NCBI backbone (NCBI:PI125735)
R:Jensen, C.H.; Krogh, T.N.; Hojrup, P.; Clausen, P.P.; Skjodt, K.; Larsson, L.I.; Englund, J. J. Biochem. 225, 83-92, 1994

A;Title: Protein structure of fetal antigen 1 (FA1). A novel circulating human epidermal pg2.
A;Reference number: S48713; MUID:95010145; PMID:7925474
A;Accession: S48713
A;Molecule type: protein
A;Residues: 24-107, D'109-282 <JEN>
R;Hojrup, P.; Jensen, C.H.; Skjodt, K.; Teisner, B.
Protein Sci. 2(Suppl.1), 259S, 1993
A;Title: Primary structure of human fetal antigen 1 (FA1), a putative homeotic glycoprotein
A;Reference number: A44549
A;Accession: A44549
A;Status: Preliminary
A;Molecule type: protein
A;Residues: 24-107, 'VASW', 112-283 <HOJ>
C;Genetics:
A;Gene: GDB:DLX1; dlx; FA1; PG2; PEF-1
A;Cross-references: GDB:9958854; OMIM:176290
A;Map position: 14q32-14q32
C;Superfamily: preadipocyte factor 1; EGF homology
C;Keywords: alternative splicing; glycoprotein; transmembrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;21-54/Domain: EGF homology #status atypical <EGF1>
F;24-383/Product: delta-like homeotic protein dlk, long splice form #status predicted <M>
F;24-282/Product: fetal antigen 1 #status experimental <MARA>
F;24-228, 302-383/Product: delta-like homeotic protein dlk, short splice form #status pre
F;57-85/Domain: EGF homology <EGF2>
F;92-124/Domain: EGF homology <EGF3>
F;131-167/Domain: EGF homology <EGF4>
F;174-205/Domain: EGF homology <EGF5>
F;212-244/Domain: EGF homology <EGF6>
F;305-327/Domain: transmembrane #status predicted <TRM>
F;94, 214/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;100/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;134/Binding site: carbohydrate (Asn) (covalent) #status absent
F;143/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;163, 251, 260/Binding site: carbohydrate (Ser) (covalent) (partial) #status experimental
F;165, 172/Binding site: carbohydrate (Asn) (covalent) (partial) #status atypical
F;222/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental

Query Match 58.8%; Score 33; DB 2; Length 383;
Best Local Similarity 62.5%; Fred. No. 97;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCLGHS 8
| | | | |
Db 220 GGTCLQHT 227

Search completed: August 25, 2004, 02:28:40
Job time : 2.80071 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 01:58:37 ; Search time 0.480427 Seconds
(without alignments)
975.448 Million cell updates/sec

Title: US-09-593-793A-113_COPY_367_375

Perfect score: 48

Sequence: 1 GATCUSHSV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	81.2	1376	1 CRBH_HUMAN	P82279 homo sapien
2	37	77.1	309	1 PNAD_HUMAN	Q96ab6 homo sapien
3	37	77.1	309	1 PNAD_MOUSE	Q64311 mus musculus
4	37	77.1	309	1 PNAD_PIG	Q28955 sus scrofa
5	35	72.9	511	1 DHAY_YEAST	P12872 saccharomyc
6	34	70.8	110	1 TNG2_HUMAN	P56847 homo sapien
7	34	70.8	898	1 PAS2_SCHAM	P22848 schistocerc
8	34	70.8	2209	1 KNTC_HUMAN	P50748 homo sapien
9	33	68.8	170	1 DEF_AGR75	Q8uid1 agrobacteri
10	33	68.8	174	1 DEF_RHIME	Q92sh6 rhizobium m
11	33	68.8	175	1 DEF_BRUME	Q8ydb4 bruceella me
12	33	68.8	176	1 DEF_RHILLO	Q98d52 rhizobium l
13	33	68.8	383	1 DLK_HUMAN	P80370 homo sapien
14	33	68.8	385	1 DLK_MOUSE	Q09163 mus musculus
15	33	68.8	443	1 YEY2_YEAST	P10356 saccharomyc
16	33	68.8	637	1 PHZE_PBECL	Q51519 pseudomonas
17	33	68.8	637	1 PHZE_PBEFL	C51791 pseudomonas
18	33	68.8	1076	1 CARB_VIBCH	Q9kbh9 vibrio chol
19	33	68.8	1077	1 CARB_VIBPA	Q87f3 vibrio para
20	33	68.8	1077	1 CARB_VIEVU	Q8dm2 vibrio vuln
21	33	68.8	1210	1 BAT8_HUMAN	Q96x47 homo sapien
22	33	68.8	1263	1 BAT8_MOUSE	Q92148 mus musculus
23	32	66.7	148	1 YF93_NEIMA	Q51134 neisseria m
24	32	66.7	246	1 TRYA_RAT	P32821 rattus norv
25	32	66.7	246	1 TRYB_RAT	P32822 rattus norv
26	32	66.7	262	1 TRPC_AZOB	P26938 azospirillu
27	32	66.7	262	1 TRPC_XANAC	Q8p447 xanthomonas
28	32	66.7	270	1 TRPC_RHOSH	Q92af7 rhodobacter
29	32	66.7	277	1 TRPC_PSEPK	Q88r6 pseudomonas
30	32	66.7	277	1 TRPC_PSEPU	P20578 pseudomonas
31	32	66.7	278	1 TRPC_PSESM	Q88a03 pseudomonas
32	32	66.7	337	1 CT86_HUMAN	Q9b219 homo sapien
33	32	66.7	705	1 GUN6_DICDI	P22699 dicryosteli

ALIGNMENTS

RESULT 1

ID	CRBH_HUMAN	STANDARD;	PRT;	1376 AA.
AC	P82279;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Crumbs protein homolog 1 precursor.			
GN	CRB1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI	TaxID=9606;			
ON	[1]			
RP	SEQUENCE FROM N.A., AND VARIANTS RP12 VAL-161; TRP-250; MET-745;			
RP	CYS-764; TYR-948; THR-1041 AND PRO-1071.			
RC	TISSUE=fetal brain, and Retina;			
RX	MEDLINE=99438399; PubMed=10508521;			
RA	van den Born L.I., van Driel M.A., de Kok Y.J.M., van Soest S.,			
RA	den Hollander A.I., ten Brink J.B., van der Pol D.J.R., Payne A.M.,			
RA	Bhattacharya S.S., Kellner U., Hoyng C.B., Westerveld A.,			
RA	Brunner H.G., Bleeker-Wagemakers E.M., Deutman A.F.,			
RA	Heckenlively J.R., Creemers F.P.M., Bergen A.A.B.;			
RT	"Mutations in a human homologue of Drosophila crumbs cause retinitis			
RT	pigmentosa (RP12).";			
RT	Nat. Genet. 23:217-221 (1999).			
RP	[2]			
RP	VARIANT LCA ARG-1100, AND VARIANT RP ARG-1181.			
RP	MEDLINE=21303018; PubMed=11389483;			
RA	den Hollander A.I., Heckenlively J.R., van den Born L.I.,			
RA	van Schooneveld M.J., Blankenagel A., Rohrschneider K., Wissinger B.,			
RA	Cruysberg J.R.M., Deutman A.F., Brunner H.G., Apfelstedt-Sylla E.,			
RA	Hoyng C.B., Creemers F.P.M.;			
RT	"Leber congenital anaurosis and retinitis pigmentosa with coats-like			
RT	exudative vasculopathy are associated with mutations in the crumbs			
RT	homologue 1 (CRB1) gene.";			
RL	Am. J. Hum. Genet. 69:198-203 (2001).			
CC	-!- FUNCTION: May be involved in cell-cell interaction in neuronal			
CC	development of the retina.			
CC	-!- SUBCELLULAR LOCATION: Extracellular (Potential).			
CC	-!- TISSUE SPECIFICITY: Preferential expression in retina, also			
CC	expressed in brain and fetal brain.			
CC	-!- DISEASE: Defects in CRB1 are the cause of retinitis pigmentosa			
CC	type 12 (RP12) [MIM:600105]. RP12 is an autosomal recessive			
CC	condition which is characterized by night blindness from early			
CC	childhood and progressive visual field loss. There is general loss			
CC	of retinal pigment epithelium throughout the retina and patients			
CC	experience severe visual impairment before the age of twenty.			
CC	-!- DISEASE: Defects in CRB1 are a cause of Leber congenital anaurosis			
CC	(LCA) [MIM:204000]. LCA designates a group of autosomal recessive			
CC	retinal dystrophies that represent the most common genetic causes			
CC	of congenital visual impairment in infants and children.			
CC	-!- SIMILARITY: Contains 15 EGF-like domains.			
CC	-!- SIMILARITY: Contains 3 laminin G-like domains.			
CC	-!- DATABASE: NAME=Mutations of the CRB1 gene;			

CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci-news/crbimut.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF154671; AAF01361.1; --
 CC HSP; P08709; IBF9.
 CC Genew; HGNC:2343; CRB1.
 CC MIM; 604210; --
 CC MIM; 204000; --
 CC MIM; 600105; --
 CC GO; GO:0007267; P:cell-cell signaling; TAS.
 CC GO; GO:0007163; P:establishment and/or maintenance of cell po. . .; TAS.
 CC InterPro; IPR000152; Asx_hydroxyl_S.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR001791; Laminin_G.
 CC Pfam; PF00038; EGF_16.
 CC Pfam; PF00034; Laminin_G.
 CC PRINTS; PR00010; EGFBL00.
 CC SMART; SM00179; EGF_CA; 8.
 CC SMART; SM00282; LamG; 3.
 CC PROSITE; PS00010; ASX_HYDROXYL; 10.
 CC PROSITE; PS00022; EGF_1; 15.
 CC PROSITE; PS01186; EGF_2; 11.
 CC PROSITE; PS00026; EGF_3; 19.
 CC PROSITE; PS01187; EGF_CA; 7.
 CC PROSITE; PS00025; LAM_G_DOMAIN; 3.
 CC EGF-like domain; Glycoprotein; Repeat; Signal; Disease mutation;
 CC Retinitis pigmentosa; Vision.
 CC SIGNAL 1 25 POTENTIAL.
 CC CHAIN 26 1376 CRUMBS PROTEIN HOMOLOG 1.
 CC DOMAIN 70 108 EGF-LIKE 1.
 CC DOMAIN 110 146 EGF-LIKE 2.
 CC DOMAIN 148 184 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 186 222 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 224 260 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 304 337 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 339 395 EGF-LIKE 7.
 CC DOMAIN 441 481 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 485 670 LAMININ G-LIKE 1.
 CC DOMAIN 672 708 EGF-LIKE 9.
 CC DOMAIN 714 885 LAMININ G-LIKE 2.
 CC DOMAIN 887 923 EGF-LIKE 10.
 CC DOMAIN 950 1137 LAMININ G-LIKE 3.
 CC DOMAIN 1139 1175 EGF-LIKE 11.
 CC DOMAIN 1177 1212 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 1214 1250 EGF-LIKE 13.
 CC DOMAIN 1255 1295 EGF-LIKE 14.
 CC DOMAIN 1297 1333 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 74 85 POTENTIAL.
 CC DI-SULFID 79 96 POTENTIAL.
 CC DI-SULFID 98 107 POTENTIAL.
 CC DI-SULFID 114 125 POTENTIAL.
 CC DI-SULFID 119 134 POTENTIAL.
 CC DI-SULFID 136 145 POTENTIAL.
 CC DI-SULFID 152 163 POTENTIAL.
 CC DI-SULFID 157 172 POTENTIAL.
 CC DI-SULFID 174 183 POTENTIAL.
 CC DI-SULFID 190 201 POTENTIAL.
 CC DI-SULFID 195 210 POTENTIAL.
 CC DI-SULFID 212 221 POTENTIAL.
 CC DI-SULFID 228 239 POTENTIAL.
 CC DI-SULFID 233 248 POTENTIAL.

FT DI-SULFID 250 POTENTIAL.
 FT DI-SULFID 305 POTENTIAL.
 FT DI-SULFID 310 POTENTIAL.
 FT DI-SULFID 327 POTENTIAL.
 FT DI-SULFID 333 POTENTIAL.
 FT DI-SULFID 343 POTENTIAL.
 FT DI-SULFID 348 POTENTIAL.
 FT DI-SULFID 385 POTENTIAL.
 FT DI-SULFID 445 POTENTIAL.
 FT DI-SULFID 450 POTENTIAL.
 FT DI-SULFID 471 POTENTIAL.
 FT DI-SULFID 676 POTENTIAL.
 FT DI-SULFID 681 POTENTIAL.
 FT DI-SULFID 698 POTENTIAL.
 FT DI-SULFID 891 POTENTIAL.
 FT DI-SULFID 896 POTENTIAL.
 FT DI-SULFID 913 POTENTIAL.
 FT DI-SULFID 922 POTENTIAL.
 FT DI-SULFID 1143 POTENTIAL.
 FT DI-SULFID 1148 POTENTIAL.
 FT DI-SULFID 1165 POTENTIAL.
 FT DI-SULFID 1181 POTENTIAL.
 FT DI-SULFID 1186 POTENTIAL.
 FT DI-SULFID 1202 POTENTIAL.
 FT DI-SULFID 1218 POTENTIAL.
 FT DI-SULFID 1223 POTENTIAL.
 FT DI-SULFID 1240 POTENTIAL.
 FT DI-SULFID 1259 POTENTIAL.
 FT DI-SULFID 1268 POTENTIAL.
 FT DI-SULFID 1285 POTENTIAL.
 FT DI-SULFID 1301 POTENTIAL.
 FT DI-SULFID 1306 POTENTIAL.
 FT DI-SULFID 1321 POTENTIAL.
 FT DI-SULFID 1332 POTENTIAL.
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 427 427 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 561 561 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 657 657 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 757 757 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 871 871 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 975 975 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1000 1000 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1190 1190 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 161 161 A -> V (in RP12).
 FT VARIANT 250 250 /FTId=VAR_011641.
 FT VARIANT 745 745 C -> W (in RP12).
 FT VARIANT 745 745 T -> M (in RP12).
 FT VARIANT 764 764 R -> C (in RP12).
 FT VARIANT 948 948 /FTId=VAR_011644.
 FT VARIANT 1041 1041 C -> Y (in RP12).
 FT VARIANT 1041 1041 M -> T (in RP12).
 FT VARIANT 1071 1071 L -> P (in RP12).
 FT VARIANT 1100 1100 /FTId=VAR_011647.
 FT VARIANT 1181 1181 I -> R (in LCA).
 FT VARIANT 1181 1181 /FTId=VAR_011648.
 FT VARIANT 1181 1181 C -> R (in RP12; with coat-like exudative
 FT VARIANT 1181 1181 vasculopathy).
 FT VARIANT 1181 1181 /FTId=VAR_011649.

```

KW Hydrolase.
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 309 AA; 34546 MW; 75FE6EA52347E485 CRC64;

Query Match 77.1%; Score 37; DB 1; Length 309;
Best Local Similarity 85.7%; Pred. No. 3.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSH 7
DB 85 GATCLTH 91
|||||
|||||

RESULT 3
PNAD_MOUSE
ID PNAD_MOUSE STANDARD; PRT; 309 AA.
AC O64311;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein N-terminal asparagine amidohydrolase (EC 3.5.1.-) (Protein
DE NH2-terminal asparagine deamidase) (N-terminal asn amidase) (NTN-
DE amidase) (PNAD) (Protein NH2-terminal asparagine amidohydrolase)
DE (PNAAs).
DE
GN Mus musculus (Mouse).
OS Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C129;
RX MEDLINE=97067079; PubMed=8910481;
RA Grigoryev S., Stewart A.E., Kwon Y.T., Arfin S.M., Bradshaw R.A.,
RA Jenkins N.A., Copeland N.G., Varshavsky A.;
RT "A mouse amidase specific for N-terminal asparagine. The gene, the
RT enzyme, and their function in the N-end rule pathway.";
RL J. Biol. Chem. 271:28521-28532(1996).
CC -1- FUNCTION: Side chain deamidation of N-terminal asparagine residues
CC to aspartate. Required for the ubiquitin-dependent turnover of
CC intracellular proteins that initiate with Met-Asn. These proteins
CC are acetylated on the retained initiator methionine and can
CC subsequently be modified by the removal of N-acetyl methionine by
CC acylaminoacyl hydrolase (AAH). Conversion of the resulting N-
CC terminal asparagine to aspartate by PNAD renders the protein
CC susceptible to arginylation, polyubiquitinylation and degradation
CC as specified by the N-end rule. This enzyme does not act on
CC substrates with internal or C-terminal asparagines and does not
CC act on glutamine residues in any position.
CC -1- SUBUNIT: Monomer. (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U57692; AAC66490.1; --
CC EMBL; U57691; AAC52885.1; --
CC MGD; MGI:108471; Ncan..
CC DR GO:0005737; C:cytoplasm; IDA.
CC DR GO:0005634; C:nucleus; IDA.
CC DR GO:0008418; F:protein N-terminal asparagine amidohydrolase. .; IDA.
CC DR GO:0008344; P:adult locomotory behavior; IMP.
CC DR GO:0007613; P:memory; IMP.
CC Hydrolase.
KW
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 309 AA; 34464 MW; BBB183EDD8F2521F CRC64;

Query Match 77.1%; Score 37; DB 1; Length 309;

```

Best Local Similarity 85.7%; Pred. No. 3.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSH 7
DB 85 GATCLTH 91

RESULT 4
PNAD_PIG STANDARD; PRT; 309 AA.
AC Q2895;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein N-terminal asparagine amidohydrolase (EC 3.5.1.-) (Protein
NH2-terminal asparagine deamidase) (N-terminal Asn amidase) (NTN-
amidase) (PNAD) (Protein NH2-terminal asparagine amidohydrolase
(PNA)).
GN NTAN1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9822;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-17; 106-118; 214-227 AND
233-248.
RC TISSUE=Liver;
RX MEDLINE=95113832; PubMed=7814382;
RA Stewart A.E., Arfin S.M., Bradshaw R.A.;
RT "The sequence of porcine protein NH2-terminal asparagine
amidohydrolase. A new component of the N-end Rule pathway.";
RL J. Biol. Chem. 270:25-28(1995).
RN [2]
RP CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=94375451; PubMed=6089117;
RA Stewart A.E., Arfin S.M., Bradshaw R.A.;
RT "Protein NH2-terminal asparagine deamidase. Isolation and
characterization of a new enzyme.";
RL J. Biol. Chem. 269:23509-23517(1994).
CC -1- FUNCTION: Side chain deamidation of N-terminal asparagine residues
to aspartate. Required for the ubiquitin-dependent turnover of
intracellular proteins that initiate with Met-Asn. These proteins
are acetylated on the retained initiator methionine and can
subsequently be modified by the removal of N-acetyl methionine by
acylaminoacid hydrolase (AAH). Conversion of the resulting N-
terminal asparagine to aspartate by PNAD renders the protein
susceptible to arginylation, polyubiquitinylation and degradation
as specified by the N-end rule. This enzyme does not act on
substrates with internal or C-terminal asparagines and does not
act on glutamine residues in any position.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U17062; AAA65019.1; -.
DR PIR; A55768; A55768.
KW Hydrolase.
FT INIT MET 0
SQ SEQUENCE 309 AA; 34629 MW; C527D692823CE547 CRC64;
Query Match 77.1%; Score 37; DB 1; Length 309;
Best Local Similarity 85.7%; Pred. No. 3.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSH 7
DB 85 GATCLTH 91

RESULT 5
DHAY_YEAST STANDARD; PRT; 511 AA.
AC P32872;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldehyde dehydrogenase 2, mitochondrial precursor (EC 1.2.1.3).
GN ALD2 OR ALDH2.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Thiesen J.;
RL Thesis (1993), Heinrich-Heine University / Duesseldorf, Germany.
CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -1- PATHWAY: Ethanol utilization; second step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -1- CAUTION: Is not present in yeast genome.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z17314; GAA78962.1; -.
DR PIR; S31308; S31308.
DR HSP; P05091; ICW3.
DR SGD; L0000075; ALD2.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; algedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
FT TRANSIT 1 21 MITOCHONDRION.
FT CHAIN 22 511 ALDEHYDE DEHYDROGENASE 2.
FT NP BIND 274 279 NAD (ADP PART) (BY SIMILARITY).
FT ACT SITE 297 297 BY SIMILARITY.
FT ACT_SITE 331 331 BY SIMILARITY.
SQ SEQUENCE 511 AA; 56466 MW; 70EDAE951B34EE4A CRC64;
Query Match 72.9%; Score 35; DB 1; Length 511;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCLSH 7
DB 258 GAACLSH 264

RESULT 6
TNG2_HUMAN STANDARD; PRT; 110 AA.
AC P56877;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TNG2 protein.
GN TNG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Homini; Hominidae; Homo.

```

CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=20056259; PubMed=10588720;
CC Hallas C., Pekarsky Y., Itoyama T., Varnum J., Bichi R.,
CC Rothstein J.L., Croce C.M.;
CC "Genomic analysis of human and mouse TCL1 loci reveals a complex of
CC tightly clustered genes.";
CC Proc. Natl. Acad. Sci. U.S.A. 96:14418-14423(1999).
CC -!- TISSUE SPECIFICITY: Expressed in many tissues.
CC -!- DISEASE: ACTIVATED IN T CELL LEUKEMIAS WITH REARRANGEMENTS AT
CC 14Q32.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF195821; AAF07204.1; -.
CC MIM; 604412; -.
CC QUERY MATCH
CC 70.8%; Score 34; DB 1; Length 110;
CC Best Local Similarity 100.0%; Pred. No. 5.1;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC 3 TCSHS 8
CC 88 TCSHS 93
CC -----
CC RESULT 7
CC FAS2_SCHAM STANDARD; PRT; 898 AA.
CC ID FAS2_SCHAM STANDARD; PRT; 898 AA.
CC AC P22648;
CC DT 01-AUG-1991 (Rel. 19, Created)
CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Fasciclin II precursor (FAS II).
CC GN FAS2.
CC OS Schistocerca americana (American grasshopper).
CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
CC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
CC NCBI_TaxID=7009;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=89043938; PubMed=3187519;
CC Harrelson A.L., Goodman C.S.;
CC "Growth cone guidance in insects: fasciclin II is a member of the
CC immunoglobulin superfamily.";
CC Science 242:700-708(1988).
CC [2]
CC SEQUENCE OF 423-436.
CC RX MEDLINE=88276943; PubMed=2839842;
CC Snow P.M., Zinn K., Harrelson A.L., McAllister L., Schilling J.,
CC Bastiani M.J., Makk G., Goodman C.S.;
CC "Characterization and cloning of fasciclin I and fasciclin II
CC glycoproteins in the grasshopper.";
CC Proc. Natl. Acad. Sci. U.S.A. 85:5291-5295(1988).
CC -!- FUNCTION: NEURONAL RECOGNITION MOLECULE. INVOLVED IN A PATHWAY
CC RECOGNITION FOR AXONS DURING THE DEVELOPMENT OF NERVE FASCICLES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03789; AAA29810.1; -.
CC PIR; A40114; A40114.
CC InterPro; IPR008957; FN III-like.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_C2.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00047; IG; 5.
CC SMART; SM00060; FN3; 2.
CC SMART; SM00408; Igc2; 3.
CC PROSITE; PS50835; IG_LIKE; 5.
CC Cell adhesion; Glycoprotein; Repeat; Immunoglobulin domain;
CC Transmembrane; Signal; Neurogenesis.
CC SIGNAL 1 22
CC CHAIN 23 898
CC FT CHAIN 23 898
CC FT DOMAIN 23 764
CC FT TRANSMEM 765 782
CC FT DOMAIN 783 898
CC FT DOMAIN 31 124
CC FT DOMAIN 134 219
CC FT DOMAIN 226 316
CC FT DOMAIN 321 423
CC FT DOMAIN 428 525
CC FT DOMAIN 548 611
CC FT DOMAIN 662 730
CC FT DISULFID 48 113
CC FT DISULFID 156 203
CC FT DISULFID 248 300
CC FT DISULFID 343 407
CC FT DISULFID 450 509
CC FT CARBOHYD 35 35
CC FT CARBOHYD 51 51
CC FT CARBOHYD 149 149
CC FT CARBOHYD 192 192
CC FT CARBOHYD 297 297
CC FT CARBOHYD 328 328
CC FT CARBOHYD 447 447
CC FT CARBOHYD 457 457
CC FT CARBOHYD 580 580
CC SEQUENCE 898 AA; 99064 MW; 07989BA4F9604AC CRC64;
CC -----
CC Query Match 70.8%; Score 34; DB 1; Length 898;
CC Best Local Similarity 71.4%; Pred. No. 50;
CC Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 3 TCSHSV 9
CC DB 449 TCLAHSI 455
CC -----
CC RESULT 8
CC KNTC_HUMAN STANDARD; PRT; 2209 AA.
CC ID KNTC_HUMAN STANDARD; PRT; 2209 AA.
CC AC P50748;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Kinetochores-associated protein 1 (Rough deal homolog) (hRod) (HsROD)
CC (Rod).
CC GN KNTC1 OR KIAA0166.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=Bone marrow;
CC MEDLINE=96281124; PubMed=8724849;
CC Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
CC -----
```



```
Db 131 ATCLQHEI 138
|||||
DEF_RHIME STANDARD; PRT; 174 AA.
AC Q92SH6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR R00419 OR SMC01101.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11491430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Leclaire V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
FL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC EMBL; AL591783; CAC41856.1; ALT_INIT.
DR HAVAP; MF_00163; -; 1.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR PRODOM; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; pep_deformyl; 1.
DR ACT_SITE 137 137 BY SIMILARITY.
FT METAL 94 94 IRON (BY SIMILARITY).
FT METAL 136 136 IRON (BY SIMILARITY).
FT METAL 140 140 IRON (BY SIMILARITY).
SQ SEQUENCE 174 AA; 19476 MW; 9782A6304F12D6D8 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 174;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ATCLSHSV 9
|||||
Db 131 ATCLQHEI 138
|||||
DEF_BRUME STANDARD; PRT; 175 AA.
RESULT 11
DEF_BRUME STANDARD; PRT; 175 AA.
```

```
AC Q9YDB4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR BMEI10264 OR BRA1035.
OS Brucella melitensis, and
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bruciellaceae; Brucella.
OX NCBI_TaxID=29459, 29461;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC SPECIES=B.suis; STRAIN=1330 / Biovar 1;
RX MEDLINE=2247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kaul M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Hailing S.W., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC EMBL; AF009665; AAN53505.1; ALT_INIT.
DR EMBL; AF014596; AAN34202.1; ALT_INIT.
DR TIGR; BRA1035; -; 1.
DR HAMAP; MF_00163; -; 1.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR PRODOM; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; pep_deformyl; 1.
DR ACT_SITE 137 137 BY SIMILARITY.
FT METAL 94 94 IRON (BY SIMILARITY).
FT METAL 136 136 IRON (BY SIMILARITY).
FT METAL 140 140 IRON (BY SIMILARITY).
SQ SEQUENCE 175 AA; 19660 MW; 873A6ABC3FEFFCE2 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 175;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      2 ATCLSHSV 9
Db      131 ATCLQHEI 138

RESULT 12
DEF_RHILO
ID DEF_RHILO STANDARD; PRT; 176 AA.
AC Q98D52;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR ML4855.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099; PubMed=11214968;
RX MEDLINE=21082930;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR ENBL; AP003005; BAB51419.1; -.
DR HAMAP; MF_00163; -.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMLYASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRPFAMs; TIGR00079; Pept_deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT ACT_SITE 137..137 BY SIMILARITY.
FT METAL 94..94 IRON (BY SIMILARITY).
FT METAL 136..136 IRON (BY SIMILARITY).
FT METAL 140..140 IRON (BY SIMILARITY).
SQ SEQUENCE 176 AA; 8897DC6FBE68C045 CRC64;

Query Match 58.8%; Score 33; DB 1; Length 176;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 ATCLSHSV 9
Db      131 ATCLQHEI 138

RESULT 13

```

```

DLK_HUMAN
ID DLK_HUMAN STANDARD; PRT; 383 AA.
AC P80370; P15803;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Delta-like protein precursor (DLK) (pG2) [Contains: Fetal antigen 1
DE (FAL)].
GN DLK1 OR DLK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RX MEDLINE=93179372; PubMed=8095043;
RA Laborda J., Sausville E.A., Hoffman T., Notario V.;
RT "dlk, a putative mammalian homeotic gene differentially expressed in
RT small cell lung carcinoma and neuroendocrine tumor cell line.";
RL J. Biol. Chem. 268:3817-3820(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland, and Placenta;
RX MEDLINE=95226449; PubMed=7711066;
RA Lee Y.L., Helman L., Hoffman T., Laborda J.;
RT "dlk, pG2 and Pref-1 mRNAs encode similar proteins belonging to the
RT EGF-like superfamily. Identification of polymorphic variants of this
RT RNA.";
RL Biochim. Biophys. Acta 1261:223-232(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RX MEDLINE=90175023; PubMed=2308864;
RA Helman L.J., Sack N., Plon S., Israel M.A.;
RT "The sequence of an adrenal specific human cDNA, pG2.";
RL Nucleic Acids Res. 18:685-685(1990).
RN [4]
RP SEQUENCE OF 24-383.
RC TISSUE=Amniotic fluid;
RX MEDLINE=95010145; PubMed=7925474;
RA Jensen C.H., Krogh T.N., Hoejrup P., Clausen P.P., Skjodet K.,
RA Larsson L.-I., Engild J.J., Teisner B.;
RT "Protein structure of fetal antigen 1 (FAL). A novel circulating
RT human epidermal-growth-factor-like protein expressed in
RT neuroendocrine tumors and its relation to the gene products of dlk
RT and pG2.";
RL Eur. J. Biochem. 225:83-92(1994).
RN [5]
RP SEQUENCE OF 24-60.
RC TISSUE=Amniotic fluid;
RX MEDLINE=93273893; PubMed=8501199;
RA Jensen C.H., Teisner B., Hoejrup P., Rasmussen H.B., Madsen O.D.,
RA Nielsen B., Skjodet K.;
RT "Studies on the isolation, structural analysis and tissue localization
RT of fetal antigen 1 and its relation to a human adrenal-specific cDNA,
RT pG2.";
RL Hum. Reprod. 8:635-641(1993).
CC
CC -!- FUNCTION: May have a role in neuroendocrine differentiation.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=Long;
CC IsoId=P80370-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P80370-2; Sequence=VSP_001377;
CC -!- TISSUE SPECIFICITY: Found within the stromal cells in close
CC contact to the vascular structure of placental villi, yolk sac,
CC fetal liver, adrenal cortex and pancreas and in the beta cells of
CC the islets of Langerhans in the adult pancreas. Found also in some
CC forms of neuroendocrine lung tumor tissue.

```


DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 49.5 kDa protein in UBP3-PET122 intergenic region.
GN YER152C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313264; PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., McSedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL Nature 387:78-81(1997).
RN [2]
RP SEQUENCE OF 1-230 FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=89083497; PubMed=2849752;
RA Ohmen J.D., Kloeckner-Gruissem B., McEwen J.E.;
RT "Molecular cloning and nucleotide sequence of the nuclear PET122 gene
RT required for expression of the mitochondrial COX3 gene in S.
RT cerevisiae";
RL Nucleic Acids Res. 16:10783-10802(1988).
RN [3]
RP REVISIONS, SEQUENCE OF 1-230 FROM N.A.
RX MEDLINE=90258894; PubMed=2160592;
RA Ohmen J.D., Burke K.A., McEwen J.E.;
RT "Divergent overlapping transcripts at the PET122 locus in
RT Saccharomyces cerevisiae";
RL Mol. Cell. Biol. 10:3027-3035(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U18917; AAB64679.1; --
DR EMBL; X07558; CAA30440.1; --
DR PIR; S50655; QOBYPT.
DR GerOnline; I39230; --
DR SGD; S0000954; YER152C.
DR KW Hypothetical protein.
SQ SEQUENCE 443 AA; 49490 MW; BCA67AD3B9D7A14 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 443;
Best Local Similarity 55.6%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GATCLSHSV 9
Db 429 GAVCKSHAI 437

Search completed: August 25, 2004, 02:27:45
Job time : 2.48043 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:22:38 ; Search time 2.43416 Seconds
(without alignments)
1166.587 Million cell updates/sec

Title: US-09-593-793a-113_COPY_367_375

Perfect score: 48

Sequence: 1 GATCLSHSV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_rhiz.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	501	6 Q95KC5	Q95KC5 macaca fasc
2	48	100.0	553	4 Q96J12	Q96J12 homo sapien
3	48	100.0	553	6 Q95KI5	Q95KI5 macaca fasc
4	42	87.5	450	11 Q8K252	Q8K252 mus musculus
5	42	87.5	553	11 Q8K0H7	Q8K0H7 mus musculus
6	39	81.2	1406	4 Q8WY70	Q8WY70 homo sapien
7	38	79.2	157	13 Q98882	Q98882 brachydanio
8	38	79.2	162	13 Q91AK7	Q91AK7 brachydanio
9	38	79.2	162	13 Q91AK6	Q91AK6 brachydanio
10	37	77.1	162	13 Q91AK5	Q91AK5 brachydanio
11	37	77.1	293	11 Q8BRX4	Q8BRX4 mus musculus
12	37	77.1	310	4 Q7Z4Z0	Q7Z4Z0 homo sapien
13	37	77.1	669	5 Q41143	Q41143 caenorhabdi
14	36	75.0	135	6 Q9GL46	Q9GL46 sus scrofa
15	36	75.0	308	6 Q46370	Q46370 bos taurus
16	36	75.0	687	10 Q94H65	Q94H65 oryza sativ

Q7ZYM2 xenopus lae
Q23133 caenorhabdi
Q93A69 gamma-prote
Q8EBW3 shewanella
Q7X4Y4 rhodospiril
Q8VE71 mus musculu
Q8C4N9 strongyloce
Q8K016 mus musculu
Q7TT36 mus musculu
Q9BK33 leishmania
Q9NPH1 homo sapien
Q9P281 homo sapien
Q8P9P9 xanthomonas
Q12479 saccharomyc
Q9P280 homo sapien
Q976P7 sulfolobus
Q81P25 bacillus an
Q81CJ7 bacillus ce
Q9N364 caenorhabdi
P97031 bacillus su
Q31589 bacillus su
Q8D160 synechococc
Q9UGL4 homo sapien
Q8NAH1 homo sapien
Q7VPB0 haemophilus
Q8WRC8 plasmodium
Q7ZZZ0 brachydanio
Q86T46 homo sapien
Q86T55 homo sapien

ALIGNMENTS

RESULT 1

Q95KC5 PRELIMINARY; PRT; 501 AA.

AC Q95KC5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB062977; BAB60745.1; -.
KW Hypothetical protein.

SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Query Match 100.0%; Score 48; DB 6; Length 501;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCLSHSV 9

Db 315 GATCLSHSV 323

RESULT 2

Q96J12 PRELIMINARY; PRT; 553 AA.

ID Q96J12

AC Q96J12

DT 01-DEC-2001 (TrEMBLrel. 19, Created)


```
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RM MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; BC031381; AAH31381.1; -.
DR ENBL; AK035428; BAC29063.1; -.
DR MGJ; MGI1922082; Protein. IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0045285; F:ubiquinol-cytochrome-c reductase complex; IEA.
DR GO; GO:0008121; F:ubiquinol-cytochrome-c reductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005805; Rieseke.
DR PROSITE; PS00200; RIESEK_2; 1.
RW Hypothetical protein.
SQ SEQUENCE 553 AA; 59742 MW; 6A474FEA630B4B13 CRC64;

Query Match 87.5%; Score 42; DB 11; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATCLSHSV 9
Db 368 ATCLSHSV 375

RESULT 6
Q8WVYO PRELIMINARY; PRT; 1406 AA.
ID Q8WVYO
AC Q8WVYO;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CB1 isoform II precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99438399; PubMed=10508521;
RA den Hollander A.I., ten Brink J.B., de Kok Y.J.M., van Soest S.,
RA van den Born I.L., van Driel M.A., van de Pol D.J.R., Payne A.M.,
RA Bhattacharya S.S., Kellner U., Hoyng C.B., Westerveld A.,
RA Brunner H.G., Bleeker-Wagemakers E.M., Deutman A.F.,
RA Heckenlively J.R., Cremers F.P.M., Bergen A.A.B.;
RA "Mutations in a human homologue of Drosophila crumbs cause retinitis
RT pigmentosa (RP12).";
RL Nat. Genet. 23:217-221(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21592405; PubMed=11734541;
RA den Hollander A.I., Johnson K., de Kok Y.J.M., Klebes A.,
RA Brunner H.G., Kruet E., Cremers F.P.M.;
RA "CB1 has a cytoplasmic domain that is functionally conserved between
RT human and Drosophila.";
RL Hum. Mol. Genet. 10:2767-2773(2001).
DR ENBL; AY043325; AAL10682.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR00152; Asx_hydroxyl_S.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001791; Laminin_G.
```

```
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_16.
DR Pfam; PF00054; laminin_G; 2.
DR PRINTS; PRO0010; EGFBL00.
DR SMART; SMO0179; EGF_CA; 8.
DR SMART; SMO0282; LamG; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 10.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00022; EGF_1; 15.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS01187; EGF_CA; 7.
DR PROSITE; PS00025; LAM_G_DOMAIN; 3.
RW EGF-like domain; Signal.
FT SIGNAL 5 25 POTENTIAL.
SQ SEQUENCE 1406 AA; 154182 MW; F2D04D20FAA6E37D CRC64;

Query Match 81.2%; Score 39; DB 4; Length 1406;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCLSHS 8
Db 1226 GATCLSHS 1233

RESULT 7
Q8882 PRELIMINARY; PRT; 157 AA.
ID Q8882
AC Q8882;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Odorant receptor 1 (Fragment).
GN OR11.1 OR_ZOR1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97075163; PubMed=8917589;
RA Weth F., Nadler W., Korsching S.;
RA "Nested expression domains for odorant receptors in zebrafish
RT olfactory epithelium.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13321-13326(1996).
DR ENBL; U72683; AAB38866.1; -.
DR ZFIN; ZDB-GENE-990415-150; Or11.1.
DR GO; GO:0006021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
RW Receptor.
FT NON_TER 1 157
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17656 MW; CF5738C349E526C4 CRC64;

Query Match 79.2%; Score 38; DB 13; Length 157;
Best Local Similarity 66.7%; Pred. No. 5.7;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCLSHSV 9
Db 110 GATCLSHSV 118

RESULT 8
Q9IAK7 PRELIMINARY; PRT; 162 AA.
ID Q9IAK7
AC Q9IAK7;
```

DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Olfactory receptor (Fragment).
 GN OR11.1 OR DR1
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20183981; PubMed=10706615;
 RA Rouquier S., Blancher A., Giorgi D.;
 RT "The olfactory receptor gene repertoire in primates and mouse:
 RT Evidence for reduction of the functional fraction in primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
 DR EMBL; AF179844; AAF40409.1; -.
 DR ZFIN; ZDB-GENE-990415-190; or11.1.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR Receptor.
 KW Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 162 AA; 18156 MW; FA77B0A46EFCB7A9 CRC64;
 Query Match 79.2%; Score 38; DB 13; Length 162;
 Best Local Similarity 66.7%; Pred. No. 5.9;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GATCLSHSV 9
 Db 114 GATCVSHLI 122
 RESULT 9
 Q9IAK6 PRELIMINARY; PRT; 162 AA.
 AC Q9IAK6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Olfactory receptor (Fragment).
 GN OR11.1 OR DR3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20183981; PubMed=10706615;
 RA Rouquier S., Blancher A., Giorgi D.;
 RT "The olfactory receptor gene repertoire in primates and mouse:
 RT Evidence for reduction of the functional fraction in primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
 DR EMBL; AF179845; AAF40410.1; -.
 DR ZFIN; ZDB-GENE-990415-190; or11.1.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR Receptor.
 KW Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 162 AA; 18156 MW; FA77B0A46EFCB7A9 CRC64;
 Query Match 79.2%; Score 38; DB 13; Length 162;
 Best Local Similarity 66.7%; Pred. No. 5.9;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GATCLSHSV 9
 Db 114 GATCVSHLI 122
 RESULT 10
 Q9IAK6 PRELIMINARY; PRT; 162 AA.
 AC Q9IAK6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Olfactory receptor (Fragment).
 GN OR11.1 OR DR3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20183981; PubMed=10706615;
 RA Rouquier S., Blancher A., Giorgi D.;
 RT "The olfactory receptor gene repertoire in primates and mouse:
 RT Evidence for reduction of the functional fraction in primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
 DR EMBL; AF179846; AAF40411.1; -.
 DR ZFIN; ZDB-GENE-990415-190; or11.1.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR Receptor.
 KW Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 162 AA; 18130 MW; E1FD981D7D448C39 CRC64;
 Query Match 77.1%; Score 37; DB 13; Length 162;
 Best Local Similarity 85.7%; Pred. No. 9.5;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GATCLSHV 7
 Db 114 GATCVSH 120
 RESULT 11
 Q9BRX4 PRELIMINARY; PRT; 293 AA.
 AC Q9BRX4;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE N-terminal Asn amidase (Fragment).
 GN NTAN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,

SQ SEQUENCE 162 AA; 18122 MW; 7877B0A46EF2B949 CRC64;
 Query Match 79.2%; Score 38; DB 13; Length 162;
 Best Local Similarity 66.7%; Pred. No. 5.9;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GATCLSHSV 9
 Db 114 GATCVSHLI 122
 RESULT 10
 Q9IAK6 PRELIMINARY; PRT; 162 AA.
 AC Q9IAK6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Olfactory receptor (Fragment).
 GN OR11.1 OR DR3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20183981; PubMed=10706615;
 RA Rouquier S., Blancher A., Giorgi D.;
 RT "The olfactory receptor gene repertoire in primates and mouse:
 RT Evidence for reduction of the functional fraction in primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2870-2874(2000).
 DR EMBL; AF179846; AAF40411.1; -.
 DR ZFIN; ZDB-GENE-990415-190; or11.1.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor activity; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR Receptor.
 KW Receptor.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 162 AA; 18130 MW; E1FD981D7D448C39 CRC64;
 Query Match 77.1%; Score 37; DB 13; Length 162;
 Best Local Similarity 85.7%; Pred. No. 9.5;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GATCLSHV 7
 Db 114 GATCVSH 120
 RESULT 11
 Q9BRX4 PRELIMINARY; PRT; 293 AA.
 AC Q9BRX4;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE N-terminal Asn amidase (Fragment).
 GN NTAN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RL 60,770 full-length cDNAs.";
 DR Nature 420:563-573 (2002).
 DR EMBL; AK041112; BAC30825.1; -.
 DR MGI; 108471; Ntani.
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0008418; F:protein N-terminal asparagine amidohydrolas. . .; IDA.
 DR GO; GO:0008344; P:adult locomotory behavior; IMP.
 DR GO; GO:0007613; P:memory; IMP.
 FT NON_TER 293
 SQ SEQUENCE 293 AA; 32710 MW; 8FE4E7DA8E561661 CRC64;
 Query Match 77.1%; Score 37; DB 11; Length 293;
 Best Local Similarity 85.7%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCLSH 7
 Db 86 GATCLTH 92
 RESULT 12
 Q72420 PRELIMINARY; PRT; 310 AA.
 AC Q72420;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE N-terminal asparagine amidohydrolase.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jiang C.L., Yu L., Fan Y.X., Tu Q., Jiang J.X., Zhao S.Y.;
 RT "Cloning and expression of a novel human cDNA homology to murine N-
 RT terminal asparagine amidohydrolase (Ntani) mRNA.";
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zhao Y., Yu L., Xin Y.R., Zhang M., Chen S.Y., Zhao S.Y.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF092440; AAP97215.1; -.
 KW Hydrolase.
 SQ SEQUENCE 310 AA; 34691 MW; 44C92804CB4CA625 CRC64;
 Query Match 77.1%; Score 37; DB 4; Length 310;
 Best Local Similarity 85.7%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCLSH 7
 Db 86 GATCLTH 92
 RESULT 13
 O44143 PRELIMINARY; PRT; 669 AA.
 ID O44143;
 AC O44143;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN C44B12.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Tin-Wollam A.;
 RT "The sequence of C. elegans cosmid C44B12.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF036692; AAB88327.1; -.
 DR PIR; T32512; T32512.
 DR WormPep; C44B12.4; CE16920.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 669 AA; 76412 MW; 863406A1ED776AF0 CRC64;
 Query Match 77.1%; Score 37; DB 5; Length 669;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GATCLSHSV 9
 Db 62 GTCLSHSI 70
 RESULT 14
 Q9GL46 PRELIMINARY; PRT; 135 AA.
 ID Q9GL46;
 AC Q9GL46;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Pref-1 (Fragment).
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fahrenkrug S.C.;
 RT "Mapping of the porcine Pref-1 gene.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007208; AAG09619.1; -.
 DR HSPF; P00750; ITPG.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR006210; IEGF.
 DR Pfam; PF00008; EGF; 1.
 DR SMART; SM00181; EGF; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 KW EGF-like domain.
 FT NON_TER 135
 SQ SEQUENCE 135 AA; 14636 MW; 1AF406B03B24EA7 CRC64;
 Query Match 75.0%; Score 36; DB 6; Length 135;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GATCLSHS 8

```

Db          31 GGTCLQHS 38
|||||
RESULT 15
O46370      PRELIMINARY;          PRT;    308 AA.
ID  O46370;
DT  01-JUN-1998 (TREMBLrel. 06, Created)
DT  01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT  01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE  Preadipocyte factor-1.
GN  DLK.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Fat;
RA  Minoshima Y., Taniguchi Y., Sasaki Y.;
RJ  Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Fahrenkrug S.C., Freking B.A., Smith T.P.S.;
RT  "Genomic Organization and Chromosomal Position of the Bovine DLX
   Gene.";
RL  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB009278; BAA23733.1; -.
DR  EMBL; AF181466; AAF00926.1; -.
DR  EMBL; AF181463; AAF00926.1; JOINED.
DR  EMBL; AF181464; AAF00926.1; JOINED.
DR  EMBL; AF181465; AAF00926.1; JOINED.
DR  EMBL; AF181462; AAD56944.1; -.
DR  PIR; JC7125; JC7125.
DR  HSP; P00740; IEDM.
DR  GO; GO:0005509; F:calcium ion binding; IEA.
DR  InterPro; IPR000152; Asx_hydroxyl_S.
DR  InterPro; IPR000742; EGF_2.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR001438; EGF_II.
DR  InterPro; IPR006209; EGF_like.
DR  Pfam; PF00008; EGF; 4.
DR  PRINTS; PR00010; EGFBLD.
DR  SMART; SM00179; EGF_CA; 1.
DR  PROSITE; PS00010; ASX_HYDROXYL; 1.
DR  PROSITE; PS00022; EGF_1; 4.
DR  PROSITE; PS01186; EGF_2; 5.
KW  EGF-like domain.
SQ  SEQUENCE 308 AA; 33005 MW; FF42B9103DD9D852 CRC64;

Query Match      75.0%; Score 36; DB 6; Length 308;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 GATCLSHS 8
Db          220 GGTCLQHS 227
|||||
Search completed: August 25, 2004, 02:31:21
Job time : 5.43416 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 01:54:26 ; Search time 171.214 Seconds
(without alignments)
912.596 Million cell updates/sec

Title: US-09-593-793A-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	2 AAW69385	Aaw69385 Prostate
2	2861	100.0	553	2 AAW71869	Aaw71869 Amino aci
3	2861	100.0	553	3 AAY82002	Aay82002 Human imm
4	2861	100.0	553	3 AAB28527	Aab28527 Protein e
5	2861	100.0	553	3 ABG94411	Abg94411 Human pro
6	2861	100.0	553	4 AAM01117	Aam01117 Human pro
7	2861	100.0	553	4 AAU69763	Aau69763 Human pro
8	2861	100.0	553	4 AAB74800	Aab74800 Prostate
9	2861	100.0	553	4 AAG99002	Aag99002 Human pro
10	2861	100.0	553	4 AAG82150	Aag82150 Human P50
11	2861	100.0	553	4 ABU71653	Abu71653 Prostate
12	2861	100.0	553	4 AAU04961	Aau04961 Human pro
13	2861	100.0	553	5 AAU10324	Aau10324 Human PRO
14	2861	100.0	553	5 ABB95222	Abb95222 Human Li-
15	2861	100.0	553	5 AAU82643	Aau82643 Human bre
16	2861	100.0	553	5 ABG61900	Abg61900 Prostate
17	2861	100.0	553	5 ABG76685	Abg76685 Prostate
18	2861	100.0	553	5 ABB77575	Abb77575 Human mas
19	2861	100.0	553	6 ABR54334	AbR54334 Prostate
20	2861	100.0	553	7 ADB13563	AdB13563 Human pro
21	2861	100.0	1079	4 AAB74830	Aab74830 Prostate
22	2861	100.0	1079	4 ABU71860	Abu71860 Prostate
23	2601	90.9	710	5 AAM50661	Aam50661 Thioresox
24	2596	90.7	595	4 AAM01318	Aam01318 Alpha pre
25	1696	59.3	359	4 ABU71887	Abu71887 Human pro

26 1677.5 58.6 530 5 AAM50662 Thioresox
27 1517 53.0 305 6 ABR54583 Prostate
28 1517 53.0 305 7 ADB14479 Human P50
29 1417.5 49.5 371 4 AAM01230 P553s sp1
30 1417.5 49.5 371 4 AAU69875 Human pro
31 1417.5 49.5 371 4 ABU71766 Prostate
32 1417.5 49.5 371 5 ABB95335 Human P55
33 1417.5 49.5 371 6 ABR54447 Prostate
34 1417.5 49.5 371 7 ADB14158 Human pro
35 1416 49.5 371 4 AAE01362 Human gen
36 1416 49.5 371 5 ABG64105 Human alb
37 1416 49.5 371 6 ADA57289 Human sec
38 1403.5 49.1 400 4 AAM01262 Human pro
39 1403.5 49.1 400 4 AAU69907 Human pro
40 1403.5 49.1 400 4 ABU71798 Prostate
41 1403.5 49.1 400 5 ABB95367 Human pro
42 1403.5 49.1 400 6 ABR54479 Prostate
43 1403.5 49.1 400 7 ADB14302 Human pro
44 1287 45.0 255 2 AAW85068 Protein e
45 1287 45.0 255 3 AAB29268 Human pro

ALIGNMENTS

RESULT 1

AAW69385

ID AAW69385 standard; protein; 553 AA.

AC AAW69385;

XX 25-MAR-2003 (revised)

DT 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone L1-12 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection;

XX therapy.

XX Homo sapiens.

XX WO9837418-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US003690.

XX 25-FEB-1997; 97US-00806596.

XX 01-AUG-1997; 97US-00904809.

XX 03-FEB-1998; 98US-00020747.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC;

XX WPI: 1998-480805/41.

XX N-PSDB; AAV58586.

XX Novel human prostate specific tumour protein and fragments - useful for

XX detecting and treating prostate cancers.

XX Example 1; Page 87-89; 141pp; English.

XX This sequence is encoded by a human prostate tumour specific gene, and

XX can be used in the method of the invention. The method is for detecting

XX prostate cancer comprises contacting a biological sample with an agent

XX able to bind an immunogenic portion of a prostate protein (such as this

XX protein sequence). An antibody which binds to an immunogenic portion of

XX the prostate protein, and the method can be used to detect, monitor

XX progression of, or treat prostate cancers. The antibody may also be

XX conjugated to a therapeutic agent for use in therapy of prostate cancers.

XX (Updated on 25-MAR-2003 to correct PR field.)

SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 2; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2e-271;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHKAKQLLVNLLTFTGLEVCLAAAGITYVPPLLEVGVEKEKFTWVLGIG 60
 DB 1 MVQRLWVSRLLRHKAKQLLVNLLTFTGLEVCLAAAGITYVPPLLEVGVEKEKFTWVLGIG 60

QY 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120

QY 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTOBECFLGGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 DB 181 IDWTSALAPYLGTOBECFLGGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

QY 241 CCPCRLAFPNLGAALLPRHLQCCMRPTRLRLFVAELCSWALMTFTFLFYDFVGEGL 300
 DB 241 CCPCRLAFPNLGAALLPRHLQCCMRPTRLRLFVAELCSWALMTFTFLFYDFVGEGL 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 AFVVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFVVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480

QY 481 RVVFGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
 DB 481 RVVFGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540

QY 541 VFDKSGLAKYSA 553
 DB 541 VFDKSGLAKYSA 553

RESULT 2
 AA71869
 ID AA71869 standard; protein; 553 AA.
 AC AA71869;
 DT 06-JAN-1999 (first entry)
 DE Amino acid encoded by prostate tumour clone L1-12.
 XX Prostate; cancer; tumour; vaccine; immunogen; clone.
 XX Homo sapiens.
 XX WO9837093-A2.
 XX 27-AUG-1998.
 XX 25-FEB-1998; 98WO-US003492.
 XX 25-FEB-1997; 97US-00806099.
 XX 01-AUG-1997; 97US-00904804.
 XX 09-FEB-1998; 98US-00020956.
 XX (CORI-) CORIXA CORP.
 XX

PI Xu J, Dillon DC;
 XX WPI; 1998-609886/51.
 DR N-PSDB; AAV61201.
 XX Polypeptides comprising immunogenic portions of prostate proteins - used
 PT in a vaccine for the treatment of prostate cancer.
 XX Example 1; Page 82-84; 130pp; English.
 XX The present sequence is an immunogenic portion of a prostate tumour
 CC protein. The immunogen, or the DNA encoding it, can be used as a vaccine
 CC for the treatment of prostate cancer. The immunogen was isolated from a
 CC prostate tumour cDNA library obtained by subtracting a prostate tumour
 CC cDNA expression library with a normal tissue cDNA library
 XX Sequence 553 AA;
 SQ

Query Match 100.0%; Score 2861; DB 2; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2e-271;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHKAKQLLVNLLTFTGLEVCLAAAGITYVPPLLEVGVEKEKFTWVLGIG 60
 DB 1 MVQRLWVSRLLRHKAKQLLVNLLTFTGLEVCLAAAGITYVPPLLEVGVEKEKFTWVLGIG 60

QY 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCVPLGASDHWGRYGRRRFFIWAISLGLLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVCLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTOBECFLGGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 DB 181 IDWTSALAPYLGTOBECFLGGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

QY 241 CCPCRLAFPNLGAALLPRHLQCCMRPTRLRLFVAELCSWALMTFTFLFYDFVGEGL 300
 DB 241 CCPCRLAFPNLGAALLPRHLQCCMRPTRLRLFVAELCSWALMTFTFLFYDFVGEGL 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 AFVVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFVVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480

QY 481 RVVFGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
 DB 481 RVVFGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540

QY 541 VFDKSGLAKYSA 553
 DB 541 VFDKSGLAKYSA 553

RESULT 3
 AA82002
 ID AA82002 standard; protein; 553 AA.
 AC AA82002;
 DT 13-JUN-2000 (first entry)
 XX Human immunogenic prostate tumour protein sequence SEQ ID NO:113.
 XX

KW Human, prostate cancer; diagnosis; tumour; gene therapy; detection;
 KW immunogenic; cytostatic; vaccine.
 XX Homo sapiens.
 XX WO200004149-A2.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-US015838.
 XX 14-JUL-1998; 98US-00115453.
 PR 14-JUL-1998; 98US-00116134.
 PR 23-SEP-1998; 98US-00159812.
 PR 23-SEP-1998; 98US-00159822.
 PR 15-JAN-1999; 99US-00232149.
 PR 15-JAN-1999; 99US-00232880.
 PR 09-AUG-1999; 99US-00288946.
 XX (CORI-) CORIXA CORP.
 PA Dillion DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
 XX WPI; 2000-171268/15.
 XX New polypeptide useful for treating and diagnosing prostate cancer
 PT comprises an immunogenic portion of prostate tumor protein.
 XX Claim 3; Page 138-139; 263pp; English.
 CC The present invention describes isolated polypeptides, comprising an
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
 CC and polynucleotides encoding them have cytostatic activity and can be
 CC used in vaccines and in gene therapy. The polypeptides and
 CC polynucleotides encoding them, antigen presenting cells which express the
 CC polypeptides, antibodies against the polypeptides and vaccines comprising
 CC them can be used for inhibiting the development of prostate cancer in a
 CC patient. The polypeptides can be used to generate antibodies or anti-
 CC idiotypic antibodies for passive immuno therapy. A portion of the
 CC polynucleotides encoding the polypeptides can be used as a probe or to
 CC modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AAY82000 to AAY82020 represent sequences used in the exemplification of
 CC the present invention
 XX Sequence 553 AA;
 SQ
 Query Match 100.0%; Score 2861; DB 3; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2e-271;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLEVCIAAGITYVPPDLLLEVGVBEKEFTMWLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLEVCIAAGITYVPPDLLLEVGVBEKEFTMWLGIG 60
 QY 61 PVGLVCVPLLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVGLVCVPLLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVLLDFCQCVFTPEALLSLDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVLLDFCQCVFTPEALLSLDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTALAPYLGTOSECEFLGLTLFLTCVAATLLVAEEALGPTPEAEGHSAPLSPH 240
 DB 181 IDWDTALAPYLGTOSECEFLGLTLFLTCVAATLLVAEEALGPTPEAEGHSAPLSPH 240
 QY 241 CCPCRARLAFNLGALLPRHLQCCMRPTLRFLVAELCSWMALMTFTLFTDFVGEGL 300
 DB 241 CCPCRARLAFNLGALLPRHLQCCMRPTLRFLVAELCSWMALMTFTLFTDFVGEGL 300
 QY 301 YQGVPAEPTGEARRHYDGVGRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYASVA 360
 DB 301 YQGVPAEPTGEARRHYDGVGRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYASVA 360

QY 361 APPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
 DB 361 APPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 DB 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553
 RESULT 4
 AAB28527
 ID AAB28527 standard; protein; 553 AA.
 XX AAB28527;
 AC AAB28527;
 XX 07-FEB-2001 (first entry)
 DT Protein encoded by human breast tumour cDNA clone P501S.
 DE Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
 XX Homo sapiens.
 OS WO200061756-A2.
 PN 19-OCT-2000.
 PD 10-APR-2000; 2000WO-US009688.
 XX 09-APR-1999; 99US-00288950.
 PR 02-JUL-1999; 99US-00346327.
 XX (CORI-) CORIXA CORP.
 PA Reed SG, Xu J, Dillon DC;
 PI WPI; 2000-638568/61.
 DR N-PSDB; AAC79473.
 XX A novel isolated polypeptide comprising an immunogenic portion of a
 PT breast cancer protein useful in the detection and treatment of breast
 PT cancer.
 XX Claim 2; Page 92-93; 95pp; English.
 PS The present sequence is encoded by a cDNA sequence which was isolated
 CC from a breast tumour cDNA library. It is provided in a specification
 CC relating to compounds for immunotherapy and diagnosis of breast cancer.
 CC Breast tumour antigens and the polynucleotides that encode them may be
 CC used in the production of a pharmaceutical composition to be used in the
 CC treatment of breast cancer. Proliferated T cells and incubated antigen
 CC presenting cells are also required. The polypeptides and polynucleotides
 CC may also be used to produce a vaccine
 XX Sequence 553 AA;
 SQ
 Query Match 100.0%; Score 2861; DB 3; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2e-271;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PN WO200151633-A2.
XX 19-JUL-2001.
XX 16-JAN-2001; 2001WO-US001574.
XX 14-JAN-2000; 2000US-00483672.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX WPI; 2001-425873/45.
XX New polynucleotide encoding a prostate-specific protein, for diagnosing,
PT monitoring and treating prostate cancer in a patient and for use in
PT vaccines.
XX Claim 2; Page 267-268; 543pp; English.
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC the antibodies are also used in the detection of cancer in a patient. The
CC cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC and (II) can be used in vaccines. The antibodies or (I) and (II) can also
CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate
CC volume. AAH93357 to AAH93944 and AAH0115 to AAH0138 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 28-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLLRHRKQAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEEKEFTWVLGIG 60
DB 1 MVQRLWVSRLLRHRKQAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEEKEFTWVLGIG 60
QY 61 PVILGLVCVPLLGASDHWRCYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVILGLVCVPLLGASDHWRCYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVLLDFCGQVCFPLEALLSLDFRDPDHCROAYSVYAFWISLGGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCFPLEALLSLDFRDPDHCROAYSVYAFWISLGGCLGYLLPA 180
QY 181 IDWTSALAPYLQTEECFLGLTLTLFTCTVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
DB 181 IDWTSALAPYLQTEECFLGLTLTLFTCTVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
QY 241 CCCRARLARFNIGALLPRHQCCSMPTRLRFVAELCSNMALMTFTLFTYDFVGEGL 300
DB 241 CCCRARLARFNIGALLPRHQCCSMPTRLRFVAELCSNMALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLQCAISLVFSLVMDRLVQRFGRTRAVLYASVA 360
DB 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLQCAISLVFSLVMDRLVQRFGRTRAVLYASVA 360
QY 361 AFVPAAGATCLSHVAVWTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFVPAAGATCLSHVAVWTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPFPFPALCGASACDVSVRVVVGPEPTA 480

DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPFPFPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
DB 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553
RESULT 7
AAU69763
ID AAU69763 standard; protein; 553 AA.
XX
AC AAU69763;
XX
DT 30-JAN-2002 (first entry)
XX
DE Human prostate cDNA encoded protein #3.
XX
KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS Homo sapiens.
XX
PN WO200173032-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US009919.
XX
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
XX WPI; 2001-639232/73.
XX
XX N-PSDB; AAS63557.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer.
XX
XX Claim 2; Page 269-270; 579pp; English.
XX
XX The invention relates to isolated prostate-specific polynucleotides,
CC polypeptides, fusion proteins of the polypeptides, antibodies raised
CC against the polypeptides (or antigenic epitopes derived from them) and
CC antigen-presenting cells expressing the polypeptides. The antibodies are
CC useful for detecting the presence of cancer, especially prostate cancer.
CC The polypeptides, polynucleotides and the antigen-presenting cells are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein, and for inhibiting the development of cancer especially prostate
CC cancer. Compositions comprising the polynucleotide and/or polypeptide are
CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polypeptide of the invention
XX
XX Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 4; Length 553;

Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTYVPPLLLEVGVEEKFTWVLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTYVPPLLLEVGVEEKFTWVLGIG 60

QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGNLAGLCCDPRLP 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGNLAGLCCDPRLP 120

QY 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSLAPYLGTQBECLFGLLTLLFLTCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
DB 181 IDWDTSLAPYLGTQBECLFGLLTLLFLTCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCRLAFRNLGALLPRLHQLCCMRPRTLRLFLVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCRLAFRNLGALLPRLHQLCCMRPRTLRLFLVAELCSWMLMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVGRMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVGRMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 AFVAAGATCLSHSVAVVTASAAITGFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFVAAGATCLSHSVAVVTASAAITGFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKCAPFPNGHYGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKCAPFPNGHYGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVFGRCICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
DB 481 RVVFGRCICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 8
AAB74800
ID AAB74800 standard; protein; 553 AA.
XX AC AAB74800;
XX DT 14-JUN-2001 (first entry)
XX DE Prostate tumour antigen predicted amino acid sequence for L1-12.
XX KW Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
XX KW prostate cancer; immunogenic; cytostatic; vaccine.
XX OS Homo sapiens.
XX FN W0200125272-A2.
XX PD 12-APR-2001.
XX PF 04-OCT-2000; 2000WO-US027464.
XX PR 04-OCT-1999; 99US-0157455P.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Skeiky YAW, Reed SG, Cheever MA;
XX WPI; 2001-245062/25.
XX DR N-PSDB; AAB02530.

XX PT Prostate specific protein and its encoding polynucleotide, useful for the
treatment and diagnosis of prostate cancer.
XX PS Claim 3; Page 157-158; 276pp; English.
XX CC The present invention describes an isolated polypeptide (I) comprising at
least an immunogenic portion of a prostate tumour antigen protein or its
variant. (I) have cytostatic activity and can be used in vaccine
production. (II), prostate tumour antigen polynucleotides, an antigen
presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
pharmaceutical composition containing (I) are useful for inhibiting the
development of cancer in a patient. Antibodies specific for prostate
specific proteins and oligonucleotides that hybridise to a polynucleotide
that encodes a prostate specific protein are useful for detecting the
presence or absence of a cancer or monitoring the progression of the
progression of a cancer, especially prostate cancer. AAB02422 to AAB2872,
AAB74798 to AAB74821 and AAB74830 are sequences used in the
exemplification of the present invention

XX SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTYVPPLLLEVGVEEKFTWVLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTYVPPLLLEVGVEEKFTWVLGIG 60

QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGNLAGLCCDPRLP 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGNLAGLCCDPRLP 120

QY 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSLAPYLGTQBECLFGLLTLLFLTCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
DB 181 IDWDTSLAPYLGTQBECLFGLLTLLFLTCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCRLAFRNLGALLPRLHQLCCMRPRTLRLFLVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCRLAFRNLGALLPRLHQLCCMRPRTLRLFLVAELCSWMLMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVGRMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVGRMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 AFVAAGATCLSHSVAVVTASAAITGFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFVAAGATCLSHSVAVVTASAAITGFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKCAPFPNGHYGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKCAPFPNGHYGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVFGRCICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
DB 481 RVVFGRCICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 9
AAG99002
ID AAG99002 standard; protein; 553 AA.
XX AC AAG99002;
XX

DT 25-SEP-2001 (first entry)
XX Human prostate-specific amino acid sequence L1-12/P501S.
DE Human prostate cancer; therapy; diagnosis; cat eye syndrome;
XX chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW prostate specific antigen; PSA.
KW Homo sapiens.
XX OS
XX WO200134802-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US030904.
XX 12-NOV-1999; 99US-00439313.
XX 18-NOV-1999; 99US-00443686.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
PI WPI; 2001-308785/32.
XX Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate-specific protein, useful in the diagnosis and therapy of
PT prostate cancer.
XX Claim 3; Page 167-168; 325pp; English.
XX The present invention describes an isolated polypeptide (P1) comprising
CC at least an immunogenic portion of a prostate-specific protein, or its
CC variant. Also described are polynucleotides (N1) encoding (P1) and
CC (N1) have cytotoxic activity and can be used in vaccine production. The
CC polypeptides, nucleic acids and antibodies from the present invention are
CC useful in the diagnosis and therapy of prostate cancer. Prostate specific
CC genes P704P, P712P, P774P, P775P and B305D are located in a genomic
CC region on chromosome 22q11.2 known as the Cat Eye Syndrome region.
CC Prostate specific antigen (PSA) P501S was located on chromosome 1.
CC AAH84671 to AAH85143 and AAG93000 to AAG93077 represent polynucleotide
CC and polypeptide sequences used in the exemplification of the present
XX invention
XX SQ Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLRHRKAQLLWLLTFLGLEYCLAAGITYVPPLLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLRHRKAQLLWLLTFLGLEYCLAAGITYVPPLLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPR 120
DB 61 PVLGLVCVPLLGASDHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPR 120
QY 121 ELALLILGVLLDFCGQVCTPTEALLSDFRPHDCRQAYSVYAFNWSLGGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCTPTEALLSDFRPHDCRQAYSVYAFNWSLGGCLGYLLPA 180
QY 181 IDWDTSAALYLTGTEBCLFGLTLTFLTCVAATLLVABEALGPTPEAGLSAPLS 240
DB 181 IDWDTSAALYLTGTEBCLFGLTLTFLTCVAATLLVABEALGPTPEAGLSAPLS 240
QY 241 CCPCRALFARNIGALLPRHQCCMRPTRLRFLVAELCSWVALMTFTFLYDFVGEGL 300
DB 241 CCPCRALFARNIGALLPRHQCCMRPTRLRFLVAELCSWVALMTFTFLYDFVGEGL 300
QY 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

DB 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
QY 361 APPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFNGHVGAGGSLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFNGHVGAGGSLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553
RESULT 10
AAG62150
ID AAG62150 standard; protein; 553 AA.
XX AAG62150;
XX 06-JUL-2001 (first entry)
XX Human P501S inventive antigen SEQ ID NO: 333.
DE Human; mouse; immunotherapy; cancer; leukaemia; WTI; Wilm's tumour gene;
KW chromosome 11p13; zinc finger transcription factor.
XX Homo sapiens.
XX WO200125273-A2.
XX 12-APR-2001.
XX 04-OCT-2000; 2000WO-US027465.
XX 04-OCT-1999; 99US-0157459P.
XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Xu J, Cheever MA, Reed SG;
PI WPI; 2001-328324/34.
XX Polypeptide comprising part of the Wilms Tumor gene product sequence is
PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT and cancer associated with WTI.
XX Disclosure; Page 212-213; 228pp; English.
XX The present invention describes compositions comprising peptides derived
CC from the Wilm's tumour protein WTI and methods for their use in treating
CC malignant diseases. Peptides derived from both the murine and human WTI
CC proteins are provided. The human WTI gene is found on chromosome 11p13,
CC and the protein was shown to be a zinc finger transcription factor. The
CC immunogenic peptides of the invention are particularly useful in the
CC diagnosis and treatment of cancer and leukaemia. The present sequence is
CC a polypeptide described in the exemplification of the invention
XX SQ Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLRHRKAQLLWLLTFLGLEYCLAAGITYVPPLLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLRHRKAQLLWLLTFLGLEYCLAAGITYVPPLLLEVGVEEKFTMWLGIG 60

QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTSLAPYLGTQEBCLFGLTLTLFTCTVAATLVABEAAALGPTPEAGLSAPLSLPH 240
 Db 181 IDWDTSLAPYLGTQEBCLFGLTLTLFTCTVAATLVABEAAALGPTPEAGLSAPLSLPH 240
 QY 241 CCPCRAFLAFNLAGLLPRLHQLCCMRPTRLRLFVAELCSMMALMTFTLFYTFDFVGEGL 300
 Db 241 CCPCRAFLAFNLAGLLPRLHQLCCMRPTRLRLFVAELCSMMALMTFTLFYTFDFVGEGL 300
 QY 301 YQGVPRAPGTEARRHYDEGVRSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGVRSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIWQLSQSVTAVMWSAAGLGLVAIFYATQ 540
 Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIWQLSQSVTAVMWSAAGLGLVAIFYATQ 540
 QY 541 VFDKSLAKYSA 553
 Db 541 VFDKSLAKYSA 553

RESULT 11

ID ABU71653 standard; protein; 553 AA.

XX AC ABU71653;

XX 10-JUN-2003 (first entry)

XX Prostate cancer specific antigen P501S.

XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
 KW immunogen; cancer; prostate specific antigen.

XX Homo sapiens.

XX Synthetic.

XX US2002192763-A1.

XX 19-DEC-2002.

XX 29-JUN-2001; 2001US-00895793.

XX 04-OCT-1999; 99US-0157455P.

XX 04-OCT-2000; 2000US-00679272.

XX 28-MAR-2001; 2001US-00822827.

XX (XUJJ/) XU J.

XX (DILL/) DILLON D C.

XX (MITC/) MITCHEM J L.

XX (HARL/) HARLOCKER S L.

XX (JIAN/) JIANG Y.

XX (KALO/) KALOS M D.

XX (FANG/) FANGER G R.

XX (RETT/) RETTER M W.

XX (STOL/) STOLK J A.

XX (DAYC/) DAY C H.

PA (VEDV/) VEDVICK T S.
 PA (CART/) CARTER D.
 PA (LISX/) LI S X.
 PA (WANG/) WANG A.
 PA (SKEI/) SKEIKY Y A W.
 PA (HEPL/) HEPLER W T.
 PA (HEND/) HENDERSON R A.
 PA (HURA/) HURAL J.
 PA (MCNE/) MCNEILL P D.
 PA (HOUT/) HOUGHTON R L.
 PA (DEAS/) Y DE BASSOLS C V.
 PA (FOYT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
 PI McNeill PD, Houghton RL, Y De BassolsCV, Foy TM;
 XX WPI; 2001-245062/25.

Prostate specific protein and its encoding polynucleotide, useful for the treatment and diagnosis of prostate cancer.

Example 1; Fig 9; 85pp; English.

XX The invention describes a fusion protein comprising at least one amino acid sequence of immunogenic portions of any of the 3 sequences not defined in the specification, or sequences having at least 70 or 90 % sequence identity to any one of the 35 sequences defined in the USPTO web site, which is encoded by any of the 4 nucleotide sequences not defined in the specification. The fusion protein, composition and methods are useful for diagnosing, preventing and/or treating cancer, particularly prostate cancer. The proteins are useful as markers to indicate the presence or absence of cancer. This is the amino acid sequence of a prostate cancer specific antigen. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?DocID=US20020192763

XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 4; Length 553;

Best Local Similarity 100.0%; Pred. No. 2e-271;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHKAQLLLVNLITFGLVCLAGITYVPPLILEVGVVEKFTMWLGIG 60
 Db 1 MVQRLWVSRLLRHKAQLLLVNLITFGLVCLAGITYVPPLILEVGVVEKFTMWLGIG 60
 QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTSLAPYLGTQEBCLFGLTLTLFTCTVAATLVABEAAALGPTPEAGLSAPLSLPH 240
 Db 181 IDWDTSLAPYLGTQEBCLFGLTLTLFTCTVAATLVABEAAALGPTPEAGLSAPLSLPH 240
 QY 241 CCPCRAFLAFNLAGLLPRLHQLCCMRPTRLRLFVAELCSMMALMTFTLFYTFDFVGEGL 300
 Db 241 CCPCRAFLAFNLAGLLPRLHQLCCMRPTRLRLFVAELCSMMALMTFTLFYTFDFVGEGL 300
 QY 301 YQGVPRAPGTEARRHYDEGVRSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGVRSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
RESULT 12
AAU04961
ID AAU04961 standard; protein; 553 AA.
XX AAU04961;
XX 24-OCT-2001 (first entry)
XX Human prostate tumour protein LI-12.
XX Human; prostate tumour protein; prostate cancer.
XX Homo sapiens.
XX US6262245-B1.
XX 17-JUL-2001.
XX 25-FEB-1998; 98US-00030607.
XX 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 09-FEB-1998; 98US-00020956.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC;
XX WPI; 2001-440862/47.
DR N-PSDB; AAS10108.
XX Novel polynucleotide encoding polypeptide comprising a portion of
PT prostate tumor protein useful for inhibiting development of prostate
PT cancer or for treating prostate cancer in a patient.
XX Example 1; Col 125-127; 105pp; English.
XX The sequence is a partial prostate tumour protein, encoded by a prostate
CC tumour specific cDNA. The DNA is useful for inhibiting the development of
CC prostate cancer or for treating prostate cancer in a patient
XX Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLLRKKAQLLNVLLTFCLEVLCAAGITVYPPLLLEVGVEEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRKKAQLLNVLLTFCLEVLCAAGITVYPPLLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCVPLIGSADHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCVPLIGSADHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDPCGQVCFTPEALLSDLFDPDHCRCQAYSVAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDPCGQVCFTPEALLSDLFDPDHCRCQAYSVAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQEBCLFGLLTLIFLTCVAATLLVAEEAALGFTBPAEGLSAPLSPH 240

Db 181 IDWDTSALAPYLGTQEBCLFGLLTLIFLTCVAATLLVAEEAALGFTBPAEGLSAPLSPH 240
QY 241 CCPCEARLAFENLGCALLPRLHOLCCRMPTLRLFVAELCSWMALMTFTLFTYTDVGEGL 300
Db 241 CCPCRARLAFENLGCALLPRLHOLCCRMPTLRLFVAELCSWMALMTFTLFTYTDVGEGL 300
QY 301 YQGVPRAEPPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPRAEPPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLVHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLVHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
RESULT 13
AAU10324
ID AAU10324 standard; protein; 553 AA.
XX AAU10324;
XX 14-FEB-2002 (first entry)
XX Human PROST 03.
XX Human; PROST 03; metastasis; prostate cancer; tumour; immune response;
XX cytostatic; vaccine.
XX Homo sapiens.
XX WO200181577-A2.
XX 01-NOV-2001.
XX 26-APR-2001; 2001WO-US013323.
XX 27-APR-2000; 2000US-0200065P.
PR 20-APR-2001; 2001US-00838785.
XX (SCHD) SCHERING AG.
XX Lau T, Lin RU, Parkes D, Parry G, Schneider DW, Steinbrecher R;
PI Van Heult PT, Wu J;
XX WPI; 2002-041404/05.
DR N-PSDB; AAS14962.
XX Novel PROST 03 polypeptides and polynucleotides useful in research,
PT diagnosis and therapeutic applications, particularly for use in cancer
PT therapeutics.
XX Claim 14; Fig 2; 77pp; English.
XX The invention relates to an isolated PROST 03 polypeptide (I) and to the
CC polynucleotide (II) encoding PROST 03. Fragments of (I) were used to
CC generate antibodies (III) to PROST 03. (III) is useful for selectively
CC destroying a cell expressing (I), and for treating a disease-state
CC associated with expression of PROST 03 in a human patient. (III) is
CC useful for diagnosing metastasis associated with (I), in a subject. (I)
CC is also useful for diagnosing and treating diseases of cell proliferation

CC such as prostate cancers. (I) is also useful for generating antibodies to
CC PROST 03. (III) is useful in detecting the levels of PROST 03
CC polypeptides in cells and tissues, and in targeting drugs to primary and
CC metastatic tumours. (I) is also useful for stimulating immune response to
CC PROST 03 containing cells. (II) is useful in diagnostic assays for
CC detecting the levels of polynucleotides encoding PROST 03 in cells and
CC tissues. (II) is useful as DNA probes, as targets for antisense and
CC ribozyme therapy, and as templates for the production of antisense
CC polynucleotides. (I) and (II) are useful in research, biological,
CC clinical and therapeutic purposes. The present sequence represents the
CC amino acid sequence of human PROST 03
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 5; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRKKAQLLLVLLTFLGVCVLAAGITYVPPILLVEGVBEKFMVVLGIG 60
Db 1 MVQRLWVSRLRRKKAQLLLVLLTFLGVCVLAAGITYVPPILLVEGVBEKFMVVLGIG 60
Qy 61 PVLGLVCPVLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCPVLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Qy 121 ELALLILGVLLDFCGVCFPLEALLSDFPDHCRQAYSVYAFWISLGGCLGVLPA 180
Db 121 ELALLILGVLLDFCGVCFPLEALLSDFPDHCRQAYSVYAFWISLGGCLGVLPA 180
Qy 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVABEALGPTPEAGLSAPLSLSPH 240
Db 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVABEALGPTPEAGLSAPLSLSPH 240
Qy 241 CCPCRLARNLGALLPRLHQLCCRPRTLRLFVAELCSWALMTFTFLYDFVGEGL 300
Db 241 CCPCRLARNLGALLPRLHQLCCRPRTLRLFVAELCSWALMTFTFLYDFVGEGL 300
Qy 301 YQGVPRAEPTGTEARRHYDEGVGRMSGLFLQCAISLVFSLVMDRLVQRFGRVAVLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGVGRMSGLFLQCAISLVFSLVMDRLVQRFGRVAVLASVA 360
Qy 361 APFVAAGATCLSHVAVVTASAAITGFTPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHVAVVTASAAITGFTPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Qy 481 RVVPGRCICLDLAILDSAFLLSQVAPSLFVMSQVTSVQSVTAQVYMSAAGLVAIFYATQ 540
Db 481 RVVPGRCICLDLAILDSAFLLSQVAPSLFVMSQVTSVQSVTAQVYMSAAGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 14

ABB95222

ID ABB95222 standard; protein; 553 AA.

XX

AC ABB95222;

XX

19-JUL-2002 (first entry)

XX

Human L1-12 protein SEQ ID NO 113.

XX

Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;

XX

gene therapy.

XX

Homo sapiens.

OS

XX
PN US2002022248-A1.
XX 21-FEB-2002.
XX 12-JAN-2001; 2001US-00759143.
XX 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 10-FEB-1998; 98US-00020956.
PR 25-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00352616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443686.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 10-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
XX (XUJJ/) XU J.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (HARL/) HARLOCKER S L.
PA (JIAN/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FANG/) FANGER G R.
PA (RETT/) RETTER M W.
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS,
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX WPI; 2002-255649/30.
XX New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer.
XX
PS Claim 2; SEQ ID NO 113; 87pp; English.
XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a
CC protein described in the invention
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 5; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRKKAQLLLVLLTFLGVCVLAAGITYVPPILLVEGVBEKFMVVLGIG 60

Db 1 MVQRLWVSRLRHRKAQLLLVNLITFGLVCLAGITYVPPLLLEVGVEKEFTMWLGIG 60
QY 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLTLLSLFTIPRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLTLLSLFTIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCCQVCFPLEALLSDLFDPDHCRCQAYSVPYAPMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCCQVCFPLEALLSDLFDPDHCRCQAYSVPYAPMISLGGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
QY 241 CCPCRLAFRNIGALLPRLHQLCCMPRTLRLFLVAELCSWMALMTFTFLFYTDVFGEG 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCMPRTLRLFLVAELCSWMALMTFTFLFYTDVFGEG 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGRTRAVLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGRTRAVLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 15
AAU82643
ID AAU82643 standard; peptide; 553 AA.
XX
AC AAU82643;
XX

DT 23-APR-2002 (first entry)
XX
DE Human breast tumour polypeptide clone #3.
XX

XX Human; breast tumour polypeptide; breast cancer; cytostatic;
KW immunostimulant.
XX
OS Homo sapiens.
XX

XX WC200198339-A2.
XX

XX 27-DEC-2001.
XX

XX 12-JUN-2001; 2001WO-US019032.
XX

XX 22-JUN-2000; 2000US-00602877.
XX

XX 12-OCT-2000; 2000US-00687507.
XX

XX 06-FEB-2001; 2001US-00778381.
XX

XX (CORI-) CORIXA CORP.
XX

XX Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
XX

XX WPI; 2002-147792/19.
XX

XX N-FSDB; ABK29012.
XX

XX Polynucleotides encoding breast tumor polypeptides, useful for treating
PT breast cancer or stimulating an immune response.

XX Claim 2; Page 144-145; 150pp; English.
XX

The invention relates to polynucleotides encoding breast tumour polypeptides. The sequences are useful for treating cancer, preferably breast cancer, in a patient or for stimulating an immune response. The polynucleotides and polypeptides are also useful in the diagnosis and monitoring of breast cancer. A method for detecting the presence of a cancer in a patient, comprises obtaining a biological sample from the patient, contacting the biological sample with a binding agent that binds to a breast tumour polypeptide, detecting in the sample an amount of polypeptide that binds to the binding agent, and comparing the amount of polypeptide to a predetermined cut-off value, therefore determining the presence of a cancer in the patient. Sequences AAU82641-AAU82655 represent human breast tumour polypeptides of the invention

XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 5; Length 553;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRHRKAQLLLVNLITFGLVCLAGITYVPPLLLEVGVEKEFTMWLGIG 60
Db 1 MVQRLWVSRLRHRKAQLLLVNLITFGLVCLAGITYVPPLLLEVGVEKEFTMWLGIG 60
QY 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLTLLSLFTIPRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLTLLSLFTIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCCQVCFPLEALLSDLFDPDHCRCQAYSVPYAPMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCCQVCFPLEALLSDLFDPDHCRCQAYSVPYAPMISLGGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
QY 241 CCPCRLAFRNIGALLPRLHQLCCMPRTLRLFLVAELCSWMALMTFTFLFYTDVFGEG 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCMPRTLRLFLVAELCSWMALMTFTFLFYTDVFGEG 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGRTRAVLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGRTRAVLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

Search completed: August 25, 2004, 02:27:08
Job time : 175.214 secs

This Page Blank (uspi)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:24:18 ; Search time 162.358 Seconds
(without alignments)
1070.372 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861
Sequence: 1 MVQRLWVSRLRRHKAQLL.....AIYFATQVVFDSKLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 31425058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2861	100.0	553	US-09-745-288-101	Sequence 101, App
2	2861	100.0	553	US-09-838-785-2	Sequence 2, Appli
3	2861	100.0	553	US-09-759-143-113	Sequence 113, App
4	2861	100.0	553	US-09-780-669-113	Sequence 113, App
5	2861	100.0	553	US-09-030-606-113	Sequence 113, App
6	2861	100.0	553	US-09-822-827-113	Sequence 113, App
7	2861	100.0	553	US-09-115-453-113	Sequence 113, App
8	2861	100.0	553	US-09-832-880-113	Sequence 113, App
9	2861	100.0	553	US-09-895-793-113	Sequence 113, App
10	2861	100.0	553	US-09-895-793-113	Sequence 113, App
11	2861	100.0	553	US-10-453-919-101	Sequence 101, App
12	2861	100.0	553	US-10-012-896-113	Sequence 113, App
13	2861	100.0	553	US-10-010-940-113	Sequence 113, App
14	2861	100.0	553	US-10-144-678A-113	Sequence 113, App
15	2861	100.0	553	US-10-005-907-13	Sequence 13, Appl

16	2861	100.0	553	14	US-10-294-025-113	Sequence 113, App
17	2861	100.0	553	15	US-10-295-027-548	Sequence 548, App
18	2861	100.0	553	15	US-10-295-027-902	Sequence 902, App
19	2861	100.0	553	16	US-10-688-838-113	Sequence 113, App
20	2861	100.0	1079	9	US-09-822-827-947	Sequence 947, App
21	2861	100.0	1079	9	US-09-895-793-947	Sequence 947, App
22	2601	90.9	710	14	US-10-296-770-4	Sequence 4, Appli
23	1696	59.3	359	9	US-09-822-827-974	Sequence 974, App
24	1696	59.3	359	9	US-09-895-793-974	Sequence 974, App
25	1677.5	58.6	530	14	US-10-296-770-5	Sequence 5, Appli
26	1517	53.0	305	14	US-10-144-678A-1029	Sequence 1029, Ap
27	1517	53.0	305	14	US-10-294-025-1029	Sequence 1029, Ap
28	1417.5	49.5	371	9	US-09-759-143-708	Sequence 708, App
29	1417.5	49.5	371	9	US-09-780-669-708	Sequence 708, App
30	1417.5	49.5	371	9	US-09-822-827-708	Sequence 708, App
31	1417.5	49.5	371	9	US-09-895-793-708	Sequence 708, App
32	1417.5	49.5	371	9	US-09-895-814-708	Sequence 708, App
33	1417.5	49.5	371	13	US-10-012-896-708	Sequence 708, App
34	1417.5	49.5	371	14	US-10-144-678A-708	Sequence 708, App
35	1417.5	49.5	371	14	US-10-294-025-708	Sequence 708, App
36	1416	49.5	371	11	US-09-833-245-852	Sequence 852, App
37	1403.5	49.1	400	9	US-09-759-143-852	Sequence 852, App
38	1403.5	49.1	400	9	US-09-780-669-852	Sequence 852, App
39	1403.5	49.1	400	9	US-09-822-827-852	Sequence 852, App
40	1403.5	49.1	400	9	US-09-895-793-852	Sequence 852, App
41	1403.5	49.1	400	9	US-09-895-814-852	Sequence 852, App
42	1403.5	49.1	400	13	US-10-012-896-852	Sequence 852, App
43	1403.5	49.1	400	14	US-10-144-678A-852	Sequence 852, App
44	1403.5	49.1	400	14	US-10-294-025-852	Sequence 852, App
45	1287	45.0	255	12	US-09-841-894-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-09-745-288-101
; Sequence 101, Application US/09745288
; Patent No. US20010018059A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-745-288-101

Query Match	100.0%;	Score 2861;	DB 9;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 3.7e-245;		
Matches 553;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVQRLWVSRLRRHKAQLLNLTLFGLVCLAGITVVPPLLEVGVEEKFMTVLIG 60		
Db	1	MVQRLWVSRLRRHKAQLLNLTLFGLVCLAGITVVPPLLEVGVEEKFMTVLIG 60		
QY	61	PVLGLVCVPLIGSADHWGRYGRRRPFIALSILLSLFLI PRAGWLACLLCPDPRPL 120		
Db	61	PVLGLVCVPLIGSADHWGRYGRRRPFIALSILLSLFLI PRAGWLACLLCPDPRPL 120		
QY	121	ELAILGLVGLLDFCGQVCFPTPLBALISDLFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180		
Db	121	ELAILGLVGLLDFCGQVCFPTPLBALISDLFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180		
QY	181	IDWTSALAPYLGTQBECLFGLLLTLIFLTCVAAILLVAEEAALGPTPEAGLSAPLSPH 240		

Db 181 IDWTSALAPYLTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Qy 241 CCPCRLAFRNLCALPRHLQCCCRMPRTLRLFLVAELCSWMLMTFTFLFYTDVFGGL 300
Db 241 CCPCRLAFRNLCALPRHLQCCCRMPRTLRLFLVAELCSWMLMTFTFLFYTDVFGGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Qy 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSLAKYSA 553
Db 541 VVFDKSLAKYSA 553

RESULT 2

US-09-838-785-2
; Sequence 2, Application US/09838785
; Patent No. US2002000945A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US2002000945A1el PROST 03
; FILE REFERENCE: 5183LAUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRKKAQLLLVNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRKKAQLLLVNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
Qy 61 PVGLVGVCPVLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVGVCPVLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Qy 121 ELALLILGVLLDFCGQVCFPTLEALLSDLPDPDHCROAYSVYAPWISIGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCFPTLEALLSDLPDPDHCROAYSVYAPWISIGGCLGYLLPA 180
Qy 181 IDWTSALAPYLTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

Db 181 IDWTSALAPYLTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Qy 241 CCPCRLAFRNLCALPRHLQCCCRMPRTLRLFLVAELCSWMLMTFTFLFYTDVFGGL 300
Db 241 CCPCRLAFRNLCALPRHLQCCCRMPRTLRLFLVAELCSWMLMTFTFLFYTDVFGGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Qy 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSLAKYSA 553
Db 541 VVFDKSLAKYSA 553

RESULT 3

US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hegler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRKKAQLLLVNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRKKAQLLLVNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
Qy 61 PVGLVGVCPVLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVGVCPVLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

QY	121	ELALLILGWLLDFGQVCF	TPLEALLSDLFRDPDHC	QAYSVYAFMISLGGCLGYLLPA	180
Db	121	ELALLILGWLLDFGQVCF	TPLEALLSDLFRDPDHC	QAYSVYAFMISLGGCLGYLLPA	180
QY	181	IDWDTSAAPYLGTQBECL	FGLLTILFUTCVAATLLVAE	EAALGPTEPAEGLSAPLSPH	240
Db	181	IDWDTSAAPYLGTQBECL	FGLLTILFUTCVAATLLVAE	EAALGPTEPAEGLSAPLSPH	240
QY	241	CCPCRARLARFNLGALL	PRLHQLCCMPERTLRRUF	VAEELCSWALMTFTLFYDFVGEGL	300
Db	241	CCPCRARLARFNLGALL	PRLHQLCCMPERTLRRUF	VAEELCSWALMTFTLFYDFVGEGL	300
QY	301	YQCVPRAEPTGEARRHY	DEGVMSGLGFLQCAISLV	FSLVMDRLVQFQTRAVYLASVA	360
Db	301	YQCVPRAEPTGEARRHY	DEGVMSGLGFLQCAISLV	FSLVMDRLVQFQTRAVYLASVA	360
QY	361	APFVAAGATCLSHSVAV	TASAALTGFTFSAQLPIL	PTLASLVHREXQVFLPKYRGDTG	420
Db	361	APFVAAGATCLSHSVAV	TASAALTGFTFSAQLPIL	PTLASLVHREXQVFLPKYRGDTG	420
QY	421	ASSEDLSMTSFLPGPK	PAPFNGHVAGGSGLLPPP	PPALCGASACDVSVRVVVGSPTEA	480
Db	421	ASSEDLSMTSFLPGPK	PAPFNGHVAGGSGLLPPP	PPALCGASACDVSVRVVVGSPTEA	480
QY	481	RVPVPERGICLDAIIL	DSAFLLSOVAPSLFMSGI	VQLSQSYTAYMWSAAGLVAIFYAQ	540
Db	481	RVPVPERGICLDAIIL	DSAFLLSOVAPSLFMSGI	VQLSQSYTAYMWSAAGLVAIFYAQ	540
QY	541	VVFDKSDDLAKYSA	553		
Db	541	VVFDKSDDLAKYSA	553		

RESULT 4

US-09-780-669-113
; Sequence 113, Application US/09780669
; Patent No. US20020051977A1

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553

```

; LIFE: FRI
 ; ORGANISM: Homo sapien
 US-09-780-669-113

Query Match 100.0%; Score 2861; DB 9; Length 553;

[illegible]

RESULT 5

RESUL 3
US-09-030-606-113
; Sequence 113, Application US/09030606

```

; Patent No. US20020081580A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS
;
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/030,606
FILING DATE: 25-FEB-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.428C3

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-606-113

Query Match      100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLVCLAAAGITYVPPLLVGVGVEKEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLVCLAAAGITYVPPLLVGVGVEKEKFTMWLGIG 60
Qy 61 PVGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Qy 121 ELALLILGVGLDFCGQVCFPTPEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPTPEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRAELAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSWMAALMTFTLFTYDFVGEGL 300
Db 241 CCPCRAELAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSWMAALMTFTLFTYDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Qy 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVLIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVLIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 6
US-09-822-827-113
; Sequence 113, Application US/09922827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 113
```

```
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-822-827-113

Query Match      100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLVCLAAAGITYVPPLLVGVGVEKEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQQLLVNLLTFGLVCLAAAGITYVPPLLVGVGVEKEKFTMWLGIG 60
Qy 61 PVGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Qy 121 ELALLILGVGLDFCGQVCFPTPEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPTPEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRAELAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSWMAALMTFTLFTYDFVGEGL 300
Db 241 CCPCRAELAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSWMAALMTFTLFTYDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Qy 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVLIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVLIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 7
US-09-115-453-113
; Sequence 113, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-115-453-113

Query Match      100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

```

RESULT 8
US-09-232-880-113
; Sequence 113, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C6
; CURRENT APPLICATION NUMBER: US/09/232,880
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-880-113

```

Db	61	PVLGLVCVPLLGSASDHWGRGYGRRRPFPIWALSUGILLSLFLIPRAGWLACGLCCDPDRPL	120
Qy	121	ELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPA	180
Db	121	ELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPA	180
Qy	181	IDWDTSLAPYLGHQESCLFGLLTLILFTCVAATLIVAEAEALGPTPEAGLSAPSLSPH	240
Db	181	IDWDTSLAPYLGHQESCLFGLLTLILFTCVAATLIVAEAEALGPTPEAGLSAPSLSPH	240
Qy	241	CCPCRARLARFNLGALLPRLHLQCCRMPTLRLRFLVAELCSWMALMTFTLFYTFDVGEG	300
Db	241	CCPCRARLARFNLGALLPRLHLQCCRMPTLRLRFLVAELCSWMALMTFTLFYTFDVGEG	300
Qy	301	YQGVPRAPGTEARRHYDGVRMGSLGLFQCAISLVPSLVMDLVBORFGTRAVYLASVA	360
Db	301	YQGVPRAPGTEARRHYDGVRMGSLGLFQCAISLVPSLVMDLVBORFGTRAVYLASVA	360
Qy	361	APFVAAGATCLSHSVAVVTASALGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420
Db	361	APFVAAGATCLSHSVAVVTASALGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420
Qy	421	ASSEDSLMTSFLPGPKPGAPPNHGVGAGGSGLLPPPPALCGASACDVSVRVVVGEPT	480
Db	421	ASSEDSLMTSFLPGPKPGAPPNHGVGAGGSGLLPPPPALCGASACDVSVRVVVGEPT	480
Qy	481	RVVPGRGICLDLAILDSAFLLSVPAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIFYATQ	540
Db	481	RVVPGRGICLDLAILDSAFLLSVPAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIFYATQ	540
Qy	541	VVFEDKSLAKYSA	553
Db	541	VVFEDKSLAKYSA	553

RESULT 9
US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuguu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Steiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Panget, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-895-793-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
 Best Local Similarity 100.0%; Pred. No. 3.7e-245;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRHRKAQLLVNLLTTCGLVCLAAAGITYVPPLLLEVGVEBEKFTWVLGIG 60
 Db 1 MVQRLWVSRLRRHRKAQLLVNLLTTCGLVCLAAAGITYVPPLLLEVGVEBEKFTWVLGIG 60

Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLIILGVLLDFCGQVCFPTLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLIILGVLLDFCGQVCFPTLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWDTSAALPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 Db 181 IDWDTSAALPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGBGL 300
 Db 241 CCPCRLAFRNIGALLPRLHQCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGBGL 300

Qy 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 Db 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

Qy 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTATMVSAAGLGLVAIFYATQ 540
 Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTATMVSAAGLGLVAIFYATQ 540

Qy 541 VVFDKSDIAKYSA 553
 Db 541 VVFDKSDIAKYSA 553

RESULT 10

US-09-895-814-113
 ; Sequence 113, Application US/09895814
 ; Publication No. US20020193296A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darick
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William T.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Hural, John
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Vinals de Bassols, Carlota
 ; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C26
 ; CURRENT APPLICATION NUMBER: US/09/895.814
 ; CURRENT FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 990
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 113
 ; LENGTH: 553
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-09-895-814-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
 Best Local Similarity 100.0%; Pred. No. 3.7e-245;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRHRKAQLLVNLLTTCGLVCLAAAGITYVPPLLLEVGVEBEKFTWVLGIG 60
 Db 1 MVQRLWVSRLRRHRKAQLLVNLLTTCGLVCLAAAGITYVPPLLLEVGVEBEKFTWVLGIG 60

Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLIILGVLLDFCGQVCFPTLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLIILGVLLDFCGQVCFPTLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWDTSAALPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 Db 181 IDWDTSAALPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGBGL 300
 Db 241 CCPCRLAFRNIGALLPRLHQCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGBGL 300

Qy 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 Db 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

Qy 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTATMVSAAGLGLVAIFYATQ 540
 Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTATMVSAAGLGLVAIFYATQ 540

Qy 541 VVFDKSDIAKYSA 553
 Db 541 VVFDKSDIAKYSA 553

RESULT 11

US-10-453-919-101
 ; Sequence 101, Application US/10453919
 ; Publication No. US20040033230A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.446C7
 ; CURRENT APPLICATION NUMBER: US/10/453.919

```
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-453-919-101

Query Match      100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVBEKMTWVLGIG 60
DB      1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVBEKMTWVLGIG 60

QY      61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLTLLSLFIPRAGWLAGLCCDPRPL 120
DB      61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLTLLSLFIPRAGWLAGLCCDPRPL 120

QY      121 ELALLILGVGLDFCCGVCFPTLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB      121 ELALLILGVGLDFCCGVCFPTLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY      181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB      181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY      241 CCPCRLAFRNGLALLPRHLQCCMPRTLRLFVAELCSWMLMTFTFLFYDFVGEGL 300
DB      241 CCPCRLAFRNGLALLPRHLQCCMPRTLRLFVAELCSWMLMTFTFLFYDFVGEGL 300

QY      301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVORFTRAVYLASA 360
DB      301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVORFTRAVYLASA 360

QY      361 APVVAAGATCLSHSVAVVTASAAITGTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB      361 APVVAAGATCLSHSVAVVTASAAITGTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY      421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB      421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY      481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIVFATQ 540
DB      481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIVFATQ 540

QY      541 VVFDKSDLAKYSA 553
DB      541 VVFDKSDLAKYSA 553

RESULT 12
US-10-012-896-113
; Sequence 113, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
```

```
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Mesgher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-113

Query Match      100.0%; Score 2861; DB 13; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVBEKMTWVLGIG 60
DB      1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVBEKMTWVLGIG 60

QY      61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLTLLSLFIPRAGWLAGLCCDPRPL 120
DB      61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLTLLSLFIPRAGWLAGLCCDPRPL 120

QY      121 ELALLILGVGLDFCCGVCFPTLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB      121 ELALLILGVGLDFCCGVCFPTLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY      181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB      181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY      241 CCPCRLAFRNGLALLPRHLQCCMPRTLRLFVAELCSWMLMTFTFLFYDFVGEGL 300
DB      241 CCPCRLAFRNGLALLPRHLQCCMPRTLRLFVAELCSWMLMTFTFLFYDFVGEGL 300

QY      301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVORFTRAVYLASA 360
DB      301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVORFTRAVYLASA 360

QY      361 APVVAAGATCLSHSVAVVTASAAITGTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB      361 APVVAAGATCLSHSVAVVTASAAITGTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY      421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB      421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY      481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIVFATQ 540
DB      481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIVFATQ 540

QY      541 VVFDKSDLAKYSA 553
DB      541 VVFDKSDLAKYSA 553

RESULT 13
US-10-010-940-113
; Sequence 113, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Xu, Jianschun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Soik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-010-940-113

Query Match      100.0%; Score 2861; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLVWSRLRRKKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKEKMTWVLGIG 60
Db 1 MVQRLVWSRLRRKKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKEKMTWVLGIG 60

Qy 61 FVLGLVCVPLGASDHWGRYGRRRPFIWALSGLILLSFLIPRAGWLAGLLCPDRPL 120
Db 61 FVLGLVCVPLGASDHWGRYGRRRPFIWALSGLILLSFLIPRAGWLAGLLCPDRPL 120

Qy 121 ELALLILGVLLDFCCQVCFTEPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCCQVCFTEPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRLARLAFNIGALLPRHLQCCMRPTLRRLFVAELCSWALMTFTLFTDFVGEGL 300
Db 241 CCPCRLARLAFNIGALLPRHLQCCMRPTLRRLFVAELCSWALMTFTLFTDFVGEGL 300

Qy 301 YQGVPRAEPTCEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGTAVYLASVA 360
Db 301 YQGVPRAEPTCEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGTAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480

Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGVAIYFATQ 540

Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

```

RESULT 14

US-10-144-678A-113

; Sequence 113, Application US/10144678A

; Publication No. US20030157089A1

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jianschun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Soik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Basols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-113

```

```

Query Match      100.0%; Score 2861; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLVWSRLRRKKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKEKMTWVLGIG 60
Db 1 MVQRLVWSRLRRKKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKEKMTWVLGIG 60

Qy 61 FVLGLVCVPLGASDHWGRYGRRRPFIWALSGLILLSFLIPRAGWLAGLLCPDRPL 120
Db 61 FVLGLVCVPLGASDHWGRYGRRRPFIWALSGLILLSFLIPRAGWLAGLLCPDRPL 120

Qy 121 ELALLILGVLLDFCCQVCFTEPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCCQVCFTEPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRLARLAFNIGALLPRHLQCCMRPTLRRLFVAELCSWALMTFTLFTDFVGEGL 300
Db 241 CCPCRLARLAFNIGALLPRHLQCCMRPTLRRLFVAELCSWALMTFTLFTDFVGEGL 300

Qy 301 YQGVPRAEPTCEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGTAVYLASVA 360
Db 301 YQGVPRAEPTCEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFGTAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480

```


Qy	481	RVVPGEGICLDLIAILDSAFLLISQVAPSLFNGISIVQLSQSVTAYMVSAAGLGLVAIVPATQ	540
Db	481	RVVPGEGICLDLIAILDSAFLLISQVAPSLFNGISIVQLSQSVTAYMVSAAGLGLVAIVPATQ	540
Qy	541	VVFDDKSDLAKYSA	553
Db	541	VVFDDKSDLAKYSA	553

RESULT 15

```

US-10-005-907-13
; Sequence 13, Application US/10005907
; Publication NO. US20030156881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: NO. US20030156881Aika, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL DEGRADATION
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ. ID NOS: 13
; SOFTWARE: Patentin version 3.1
; SEQ. ID NO 13
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-005-907-13

```

Query Match	100.0%;	Score 2861;	DB 14;	Length 553;
Best Local Similarity	100.0%;	Pred. NO. 3.7e-245;		
Matches 553;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	MYQRLWVSELLHRKAOILLNNLITFGLEVCLAAGITIVPPLLLEVCUGVEKFTMWLIGT	60
Db	1	MYQRLWVSELLHRKAOILLNNLITFGLEVCLAAGITIVPPLLLEVCUGVEKFTMWLIGT	60
QY	61	PVILGLVCVPLLGSADHWRGRYGRRRPIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL	120
Db	61	PVILGLVCVPLLGSADHWRGRYGRRRPIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL	120
QY	121	ELALLILGVLLDFCGQVCFPTPEALLSDFRDPHCRQAYSVYAFMISLGGCLGVLLPA	180
Db	121	ELALLILGVLLDFCGQVCFPTPEALLSDFRDPHCRQAYSVYAFMISLGGCLGVLLPA	180
QY	181	IDWDTSAAPYLIGTQEELFGILLTILFTCVAAATLLVAEEAALGPTPEASGLSAPSLSPH	240
Db	181	IDWDTSAAPYLIGTQEELFGILLTILFTCVAAATLLVAEEAALGPTPEASGLSAPSLSPH	240
QY	241	CCPCBARIAFNLGALLPRLHQLCCRPMTILRLFVAELCSWMAALMTFTLYTDFVGBGL	300
Db	241	CCPCBARIAFNLGALLPRLHQLCCRPMTILRLFVAELCSWMAALMTFTLYTDFVGBGL	300
QY	301	YQGVPRAEPTGEARHYDEGVRMGSLGFLQCAISLVFSLVMDRLVORFCTRAVYLASVA	360
Db	301	YQGVPRAEPTGEARHYDEGVRMGSLGFLQCAISLVFSLVMDRLVORFCTRAVYLASVA	360
QY	361	APFVAAGATCLSHSVAVVTAGAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG	420
Db	361	APFVAAGATCLSHSVAVVTAGAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG	420
QY	421	ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTFA	480
Db	421	ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTFA	480
QY	481	RVPVGRGICLDLAILDSAPFLSOVAPSLFPMGSIIVQLSQSVTAYMYVSAAGLIVAIYFATQ	540
Db	481	RVPVGRGICLDLAILDSAPFLSQVAPSLFPMGSIIVQLSQSVTAYMYVSAAGLIVAIYFATQ	540
QY	541	VVFVKSDSLAKYA	553

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:00:53 ; Search time 49.1993 Seconds
(without alignments)
1081.193 Million cell updates/sec

Title: US-09-593-793A-113
Perfect score: 2861
Sequence: 1 MYQRLWVSRLLRHRAQLLL.....AIYFATQVFDKSLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	347.5	12.1	515	2 T14340	sucrose-proton tra
2	337.5	11.8	523	2 T12198	sucrose transport
3	330	11.5	516	2 JQ2389	sucrose transport
4	328.5	11.5	594	2 G84441	probable sucrose-p
5	324.5	11.3	525	2 S28052	sucrose transport
6	323.5	11.3	512	2 S38196	sucrose transport
7	323.5	11.3	512	2 G86360	probable sucrose-p
8	317.5	11.1	507	2 S48789	sucrose transport
9	311	10.9	533	2 S43142	sucrose transport
10	305.5	10.7	501	2 T14339	sucrose-proton tra
11	297.5	10.4	428	2 S48788	sucrose transport
12	294	10.3	510	2 S38657	sucrose transport
13	294	10.3	523	2 S51114	sucrose-proton sym
14	292.5	10.2	512	2 P96741	probable sucrose t
15	292	10.2	537	2 T02982	probable sucrose t
16	286	10.0	474	2 A86234	hypothetical prote
17	281.5	9.8	491	2 E96891	probable sucrose-p
18	279.5	9.8	513	2 S38197	sucrose transport
19	276.5	9.7	492	2 A84520	probable sucrose-p
20	273	9.5	503	2 S52377	sucrose transport
21	235.5	8.2	553	2 T38541	probable sucrose c
22	191	6.7	452	2 F75217	hypothetical prote
23	187.5	6.6	541	2 B87532	transporter, proba
24	142	5.0	544	2 S75896	melibiose carrier
25	139.5	4.9	454	2 A75444	hypothetical prote
26	138.5	4.8	430	2 E75217	transporter PAB217
27	136	4.8	418	2 B87536	membrane protein,
28	133	4.6	389	2 G83413	probable MPS trans
29	128.5	4.5	422	2 G83503	probable MPS trans

30 128.5 4.5 594 2 A83096 probable permease
31 127.5 4.5 472 2 F82639 resistance protein
32 127 4.4 451 2 E81781 probable integral
33 126 4.4 394 2 AB3073 tetracycline resis
34 126 4.4 394 2 G98213 drug efflux protei
35 126 4.4 451 2 A81206 sugar transporter,
36 125.5 4.4 399 2 JQ1479 tetracycline resis
37 125 4.4 391 2 S74888 hypotheical prote
38 125 4.4 407 2 H75515 tetracycline-efflu
39 124.5 4.4 399 1 S38656 teta protein - Pse
40 123.5 4.3 400 2 AG2866 MFS permease (drug
41 123.5 4.3 400 2 D97643 probable efflux pr
42 123 4.3 440 2 AD0986 hypotheical metab
43 121.5 4.2 398 2 C90349 multidrug-efflux t
44 121.5 4.2 503 2 S43017 puromycin resistan
45 121 4.2 483 2 AF2363 hypotheical prote

ALIGNMENTS

RESULT 1

T14340
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14340
R:Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot
A:Reference number: Z17991; MUID:99063785; PMID:9847123
A:Accession: T14340
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-515 <SHA>
A:Cross-references: EMBL:Y16768; NID:g29698893; PIDN:CAA76369.1; PID:g29698894
A:Experimental source: cultivar Nantaise; root
C:Genetics:
A:Note: SUT2
C:Superfamily: common tobacco sucrose transport protein

Query Match 12.1%; Score 347.5; DB 2; Length 515;
Best Local Similarity 25.9%; Pred. No. 1e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
QY 17 QLILVLLTFCLEVCVLAAGITYVPPLLEVGVEEKFTWVLGIGFVLGVCVPLGASD 76
Db 34 KLVLVAAIAGVQFGWALQSLTPTVYQLLGI PKWAAIWLCOPIGMLVQPIGVYSD 93
QY 77 HWRGRRRRPFIWALSIGLILSLFLIPRAGWLGL-----LCPDRPLELALLIGV 129
Db 94 HCQSSFGRRRPFIAAGACVAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF 150
QY 130 GLILFCGVCVTPLEALISDLFR-DPDHCRQAYSVAFMISLGGCLGV-----LL 178
Db 151 WILDVANMLGCGPCRALLDLCSGDTMRMSANAFYFFMVAWMLGVAAGSYNNLYKLF 210
QY 179 PAIDWDTSALAPYLGTOECLFGLITLFTCVATLLVAEEAALGTEPAEGLSAPLS 238
Db 211 PFS--KTHACDLYCANLKSCFIISIALIITTVALSVRENS--GPPDDADAAEEP--- 263
QY 239 PHCPCPARLARFNLGALLPRLHOLCCMPRTLRLFVAELCSWMLMTFLFTDFVGE 298
Db 264 ----PSSGKIPEV--FGELLGALKDL----PRMILLIIVTCLNNIAWFPFFLFTDMMGR 313
QY 299 GLYQGVPAEPTGTEARRHYDEGRVMSGLGFLOCAISLVFSLVMDRLVQRFGRVAVLAS 358
Db 314 EYGGT--AGG-----KLYDQGVKAGALGLLNSVWGLTSIAVEYLVKRGVGVKI-LWG 366
QY 359 VAAPFVAAGATCL-----SHSV-----AVVTASA-----ALTGFTSALQIL 395
Db 367 FWNFTLAIGLVMTVVVSVKVAQCHQHSANGQLPPSAGVKAGALSLSILGILPISITYSI 426

```

Qy      396   PYTLASLYHREKVFPLPKYRGDTGCGASSEDSIMTSFLPCKPG-----APPNGHVCA 448
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      427   PFLASIYSGGAGGGLSLGVNLINAIIVVQMIVSLAGPFDSLFGGNNLPFAFVGATSA 486
           ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy      449   GSGSLLP----PPPALCGASACDVS 469
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      487   AIGVLIAIALLPKPKDAASKLSL 511
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 2
Tl2198
sucrose transport protein - fava bean
C;Species: Vicia faba (fava bean)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Jun-2000
C;Accession: Tl2198
R;Weber, H.; Borisjuk, L.; Heim, U.; Sauer, N.; Wobus, U.
Plant Cell 9, 895-908, 1997
A;Title: A role for sugar transporters during seed development: molecular charac
A;Reference number: Z17451; MUID:973555984; PMID:9212465
A;Accession: Tl2198
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-523 <WEB>
A;Cross-references: EMBL:Z93774
C;Genetics:
A;Gene: sut
C;Superfamily: common tobacco sucrose transport protein

Query Match          11.8%; Score 337.5; DB 2; Length 523;
Best Local Similarity 25.8%; Pred.No. 5.8e-18;
Matches 134; Conservative 76; Mismatches 185; Indels 125; Gaps 20
```

C/Species: Solanum tuberosum (potato)
C/Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 21-Jul-2000
C/Accession: JQ2389; S40310
C/Riesmeier, J.W.; Hirner, B.; Frommer, W.B.
Plant Cell 5, 1591-1598, 1993
A/Title: Potato sucrose transporter expression in minor veins indicates a role in phloem
A/Reference number: JQ2389; MUID:94146554; PMID:8312741
A/Accession: JQ2389
A/Molecule type: mRNA
A/Residues: 1-516 <R1>
A/Cross-references: EMBL:X69165; NID:G439293; PIDN:CAA48915.1; PID:G439294
A/Experimental source: cv. Desiree
C/Comment: The gene encoding for this protein is highly expressed in mature leaves.
C/Superfamily: common tobacco sucrose transport protein
C/Keywords: Glycoprotein; transmembrane protein
F/31-53/Domain: transmembrane #status predicted <TM1>
F/67-86/Domain: transmembrane #status predicted <TM2>
F/103-122/Domain: transmembrane #status predicted <TM3>
F/141-160/Domain: transmembrane #status predicted <TM4>
F/180-200/Domain: transmembrane #status predicted <TM5>
F/226-248/Domain: transmembrane #status predicted <TM6>
F/285-304/Domain: transmembrane #status predicted <TM7>
F/331-349/Domain: transmembrane #status predicted <TM8>
F/366-385/Domain: transmembrane #status predicted <TM9>
F/409-427/Domain: transmembrane #status predicted <TM10>
F/429-448/Domain: transmembrane #status predicted <TM11>
F/3-92/Binding site: carbohydrate (Aen) (covalent) #status predicted

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C/Accession: G84441
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koc, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss; D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: G84441
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-594 <STO>
A/Cross-references: GB:AE002093; NID:G3461813; PIDN:AAC32907.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g02860
A/Map position: 2
C/Superfamily: common tobacco sucrose transport protein

Query Match 11.5%; Score 328.5; DB 2; Length 594;
Best Local Similarity 24.1%; Pred. No. 3.2e-17;
Matches 120; Conservative 72; Mismatches 174; Indels 131; Gaps 18;

Qy 22 NLTFGLVCLAGITY-----VPLLLEVGVEEKMTMVLGIGPVGLVLCVPLGSG 73
Db 59 SLTVLVSCTVAAAGVQFGWALQSLTPTIQTIGSHAFSSFTLWCGPTGLWQVFPVGI 118
Qy 74 ASHWGRYGRRRPPTWALSGLLSLELIPRAGWLA--GLLCPD-----PRPLE 121
Db 119 WSKCTSKYGRRRPPTWALSGLLSLELIPRAGWLA--GLLCPD-----PRPLE 175
Qy 122 LALLILGVGLDPCGVCFTEPLBALLSDLRDPDCHCRQYSAVAFMISGGCLGYLLPAI 181
Db 176 AVFPIIGFWLLDANNTVQGFARALLADL--SGPDQNTANAVFCLMWAIGNILGFSAGAS 234
Qy 182 ----DWDTSALAPYLGTQECLEF-----LTLFELT-CVAATLLVAEEAALGPTTP 228
Db 235 GKQEW-----PFFL-TSRACCAACGNLKAFLAVVFLTCTLVITYFAKEIPTTSNKP 288
Qy 229 AE-GLGAPSLSPHCCPCRRALAPRNL-----LQSKGLEHKLNNGTANGIKYVERDTEQFGNSNEHQ 253
Db 289 TRICDSAPLDD-----LQSKGLEHKLNNGTANGIKYVERDTEQFGNSNEHQ 340
Qy 254 -----GALLPRLHQLCCMRPTLRLFVABLCSWMAIMTFLTYTFVGEGLYQGV 305
Db 341 DETYVDGPGSVLVNLLTSLAHLPPAMHSVLIVMALTWLSFPFFLEDTDMMCREVYHGD 400
Qy 306 RAEPGTEARHDEYGVWMSGLFLQCAISLVFSLVMDRLVQRFGRVY--LASVAAPV 364
Db 401 TGD--SLHMLYDQGVREGALGLLNSVLGSSFLIEPCQRMGRVWALSNTFTVPAC 458
Qy 365 AAGATCLSL-----HSVAVVTASALTOFTFSALQILPYTLASLY 403
Db 459 MAGTAVISLMSLSDKNGIEYIMKGNBETTRTAIV--FALLGFLAITVSVFSGVTA-- 514
Qy 404 HREKQVFLPKVGRDTGG 420
Db 515 -----EVTADSG 522

RESULT 5
S28052
sucrose transport protein - spinach
N/Alternate names: sucrose carrier protein; sucrose permease
C/Species: Spinacia oleracea (spinach)
C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C/Accession: S28052
R/Riesmeier, J.W.; Willmitzer, L.; Frommer, W.B.
EMBO J. 11, 4705-4713, 1992
A/Title: Isolation and characterization of a sucrose carrier cDNA from spinach by functi
A/Reference number: S28052; MUID:93099843; PMID:1464305
A/Accession: S28052
A/Molecule type: mRNA
A/Residues: 1-525 <RIE>

A/Cross-references: EMBL:X67125; NID:G21318; PIDN:CAA47604.1; PID:G21319
C/Superfamily: common tobacco sucrose transport protein
C/Keywords: transmembrane protein

Query Match 11.3%; Score 324.5; DB 2; Length 525;
Best Local Similarity 23.2%; Pred. No. 5.7e-17;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

Qy 23 LLTFGLSVCLAGITY-----VPLLLEVGVEEKMTMVLGIGPVGLVLCVPLGSGA 74
Db 35 LKGLGLVASAAGVQFGWALQSLTPTIQTIGSHAFSSFTLWCGPTGLWQVFPVGI 94
Qy 75 SDHWGRYGRRRPPTWALSGLLSLELIPRAGWLA--GLLCPD-----RPLELALLIL 127
Db 95 SDRCTSRFGRRPPTIAGAALVAVAGLI--GPAADIGAASGDPGTGNVAKPRAIAPV 151
Qy 128 GVLGLDFCGQVCFTEPLBALLSDLR--DPDCHCRQYSAVAFMISGGCLGYLLPAID---- 182
Db 152 GFWILDVANNTLQGPCRALLADMAAGSOTKTRYANAFSFFMALGNIGGYAAGSYSLYT 211
Qy 183 ----WDTLSALAPYLGTQECLEFGLLT-LIFLTCVAAATLLVAEEAALGPTTPAEGLSAPSL 237
Db 212 VFPTKTAACDVYCANLKSCEFFISITLILVILALSVKREQITIDEIQEEDLKRN 271
Qy 238 SPHCCPCRRALAPRNLGALLPRLHQLCCMRPTLRLFVABLCSWMAIMTFLTYTFV 297
Db 272 SSGC-----ARLPF--FGOLIGALKDL-----PKPMLILLVLTALNWIAMFFELFDTDMWG 321
Qy 298 EGLYGVPRAPGTEARHDEYGVWMSGLFLQCAISLVFSLVMDRLVQRFGR--TRAVY 355
Db 322 KEVYGGT-----VGEKLYDQGVHAGALGMINSVLVGWSLSIEGLARWVGAKRLWG 375
Qy 356 LASVAAPFVAAGATCLSHSVAVVTASALTOFTFSALQILPYTLASLYHREKQVFLPKYR 415
Db 376 IVNII-----AVCLAMTV-LVTKSAE-----HFRDSHIM----- 405
Qy 416 GDTGGAESSELSMTSFLPGKPGFNGHVGAGSGGLPPPPALCGASACDVSVRVVVG 475
Db 406 -----GSAVPPPPPPA--GVKGGALAIFAVLG 429
Qy 476 EPTEARV-VP-----GRGICLDLAILDSAFLLSQV-----AP 506
Db 430 IPLAITFPIPALASIFSASSSGGSLGVNLAIIVVPQMFVSVTSVSGPDMFGGGNLP 489
Qy 507 SLFMGSIVQLSQSVTAY 523
Db 490 AFVGVAAVATASAVLSF 506

RESULT 6
S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N/Alternate names: sucrose-proton symporter SUC2
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C/Accession: S38196; T00773
R/Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A/Description: SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana.
A/Reference number: S38196
A/Accession: S38196
A/Molecule type: mRNA
A/Residues: 1-512 <SAU>
A/Cross-references: EMBL:X75382; NID:G407091; PID:G407092
R/Yosotskaja, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.; Li
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel,
submitted to the EMBL Data Library, July 1998
A/Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequence
A/Reference number: Z14202
A/Accession: T00773
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-490, 'A', 492-512 <VYS>


```

Db      132 PRAIAVVFVGFVWLVANNMLOQPCFALLADLSGGKARVRTSNAPFSFMAVGNVGYAA 191
QY      179 PAID-----WDTGALAPYLGTOECLF-GLLTILFLTCVAAILLVAEBAALGTEPA 229
Db      192 GYSRLCKIPFPTKACDIYCANLKSCFFIAVFLLSLITALT--VYRENELPEKDEH 249
QY      230 EGLSAPSLPHCCPCRRALAFRNLAGALLPRLHQLCCRMPTURRLFVABELCSWALMTFT 289
Db      250 E-----IDEKAGRSKVF--FGEIFGALKDL-----PRPMWILLVTSLNWIARFFF 297
QY      290 LFTVDVFGSLGVQVPRAPGTEARRHYDEGVMSGLGFLQCAISLVFSLVMDRLVORF 349
Db      298 LYDTDMKEVYGG--KVGDG-----RLYDLGVHAGALGLLNSVVLGFMSLSVVEFLGRKI 351
QY      350 GREAVVLASVAAPVVAAGATCLSHSVAVV-----TASAALTGTFFSALQI-- 394
Db      352 G-GVKRLWGLNFVL---AVCWALIVLVTKMAEKSRQYDAHGTLMAPTSGVIGALTLP 407
QY      395 -----LPYTLASLYHREK-----QVFLPKYRGDTGASSEDLSMTS 430
Db      408 VLGIPLAVTFSPFALASIFSSNAGSGQGLSLGVNLAIWVFMQLVSIAGGPDWDLFGG 467
QY      431 FLBPFGAPFPPNGHVAGGSG-----LPPPPA 459
Db      468 NLPG-----FIVGAVAAASGILALTMLPSPPA 495

```

RESULT 9

S43142

sucrose transport protein - castor bean

N;Alternate names: sucrose carrier

C;Species: Ricinus communis (castor bean)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000

C;Accession: S43142

R;Weig, A.; Komor, E.

submitted to the EMBL Data Library, March 1994

A;Description: A sucrose carrier from Ricinus communis.

A;Reference number: S43142

A;Accession: S43142

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-533 <WEI>

A;Cross-references: EMBL:Z31561; NID:g468561; PID:g468562

C;Superfamily: common tobacco sucrose transport protein

```

Query Match      10.9%; Score 311; DB 2; Length 533;
Best Local Similarity 23.3%; Pred. No. 6.1e-16;
Matches 134; Conservative 95; Mismatches 225; Indels 120; Gaps 19;

```

```

QY      17 QLLVNLTLFGLVCLAGITVPPILLLEVGVEEKFWTWLIGIPVLGVCVPLLGASD 76
Db      37 KVVNVASIAAGIQFGWALQSLTTPYQULGIPHTWAAFIWLCGPISGMVLPQIVGYHSD 96
QY      77 HWRGRRRRPPTIWAISLGLLSLFLIPRA---GWLAG-LLCPDPRPLEALLILGVGLL 132
Db      97 RCTSRGRRRPPTIAGNAFAVAIFLIGVADLGHLSGDSLDKSPKTRALIAFVVGFWIL 156
QY      133 DFCQGVCTTPLEALLSDLF-RDPDHCQAVSVAFMISLGGCLGYLLPAID-----W 183
Db      157 DVANNMLQGPCRALLADLSGTSOKTRTANALSFPMVAVGNVLGYAAGAYTHLYKLFPFT 216
QY      184 DTSALAPYLGTOECLFGLLTILFLTCVAATLVAEFAALGPTPEAGL-----SAP 235
Db      217 KTTACDVYCANLKSCFFISVLLSLTLVLSVVEK-----PWSPDQAVDNEDDTASQA 272
QY      236 SLSPHCCPCRRAL--AFRNIGALLPRLHQLCCRMPTLRLFVAELCSWALMTFTFLVYT 293
Db      273 SSSAQPMFPFGEILGAFKNL-----KRPMTILLIVTCLNIAWPPFLFPT 318
QY      294 DFVCEGLYQGVPRAPGTEARRHYDEGVMSGLGFLQCAISLVFSLVMDRLVORFQTRA 353
Db      319 DWMGREVYGG--DSSGSAEOLKLYDRGRAGALGMLNSVVLGFTSLGVEVLARGVG-GV 375

```

```

QY      354 VYLASVAAPVVAAGATCLSHSVAVVTASAALTGTFFSALQILPYTLASLYHREKQVLPK 413
Db      376 KRLMGIVNFVL---AVCLAMTV-LVTQAESE-----RR 405
QY      414 YRGDTGASSEDLSMTSFLPGPKGAPFPPNGHVAGGSGLLPPPPALCGASACDVSVRVV 473
Db      406 FATVSGGAK-----VP-----LPPP---SGVKAGALALFAV 433
QY      474 VGEPTEARV-VP-----GRGICLDLAILDSAFLLSOVAPSLFMGSIVQL--SQS 519
Db      434 MGVFQALTYSIIPFALASIFNTSGAGQGLSLGVNLNLSIVIPQMLVSVAAAGPMDALFOG 493
QY      520 VTATVWSAAGLGLVAIYFATQVVFQKSDLAKYSA 553
Db      494 LPAFVVGAVAAALASGIFALTMLPSPQPDMPSAKA 527

```

RESULT 10

T14339

sucrose-proton transport protein - carrot

N;Alternate names: sucrose/H⁺ symporter protein

C;Species: Daucus carota (carrot)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C;Accession: T14339

R;Shakya, R.; Sturm, A.

Plant Physiol. 118, 1473-1480, 1998

A;Title: Characterization of source- and sink-specific sucrose/H⁺ symporters from carrot

A;Reference number: 217991; MUID:99063785; PMID:9847123

A;Accession: T14339

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-501 <SHA>

A;Cross-references: EMBL:Y16766; NID:g2969886; PIDN:CAA76367.1; PID:g2969887

A;Experimental source: cultivar Nantaise; leaf

C;Genetics:

A;Gene: SUR1a

C;Superfamily: common tobacco sucrose transport protein

```

Query Match      10.7%; Score 305.5; DB 2; Length 501;
Best Local Similarity 21.6%; Pred. No. 1.5e-15;
Matches 124; Conservative 91; Mismatches 213; Indels 145; Gaps 14;

```

```

QY      18 LLLVNLTLFGLVCLAGITVPPILLLEVGVEEKFWTWLIGIPVLGVCVPLLGASDH 77
Db      30 LLRVASVACIQFGWALQSLTTPYQELGIPHAWSIIWLCGPLSGLLVCPVIGHMSDQ 89
QY      78 WRGRRRRPPTIWAISLGLLSLFLIPRAGWLAGLL--CPDPRPLEALLILGVGLLDFC 135
Db      90 CTSKYGRRRPPIVAGGTAILLAVIIIAHSADIGLLGDTADNKTWIAVAFVIGFWILDVA 149
QY      136 QGVCTPLEALLSDLF-RDPDHCQAVSVAFMISLGGCLGYLLPAIDWDTSALAPYLG 194
Db      150 NMTQGPCRALLADLTGNDARRTRVANAYPSLFWAIGNVLGY-----ATGAYSGW 199
QY      195 QECFLGLTLTLFTLC-----VAATLLVAEEAAL-----GPTPEAEG 232
Db      200 YKVPFSLTSSCTINCANKLSAFYIDIFIITITVISIAKERPRISSQDGPQFSEDGT 259
QY      233 SAPLSLPHCCPCRRALAFRNIGALLPRLHQLCCRMPTLRLFVAELCSWALMTFTFLFY 292
Db      260 AQ---SGHI-----EEAFWLFLGTFRLLPGSVVWILLVTCNLNMGWFFFLFD 305
QY      293 TDFVCEGLYQGVPRAPGTEARRHYDEGVMSGLGFLQCAISLVFSLVMDRLVORFQTR 352
Db      306 TDMGREIYGEPE-----NOGQSYSDGVKMGAFGLMKNVSVLGITSVLMEKLCRIWGS 359
QY      353 AVYLASVAAPVVAAGATCLSHSVAVVTASAALTGTFFSALQILPYTLASLYHREKQVLP 412
Db      360 FWM-----GLSNILMTICPFAMLLITFIKQNDY----- 388
QY      413 KYRGDTGASSEDLSMTSFLPGPKGAPFPPNGHVAGGSGLLPPPPALCGASACDVSVRV 472

```

Db 389 -----CTNPPN-----GIVISALIVFA 406
 Qy 473 VVEPTARV-VP-----GKICLDAILDSAFLLSQVAPSLFVMSIVQL----- 516
 Db 407 ILGIPLAITYSVFVALVSTRIESLGLQGLQSMGLVNLAIWVQVIVSLGSGPMDQLFGGG 466
 Qy 517 -SOSVTAVMYSAAGLGVAVIYFATVGVFDSKL 548
 Db 467 NSPFAVVAALSAFAAGLIALIARRPRVDKSL 499

RESULT 11
 S48788
 sucrose transport protein - tomato (fragment)
 C/Species: Lycopersicon esculentum (tomato)
 C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jun-2000
 C/Accession: S48788
 R/Buerkle, X.Y.Z.; Frommer, W.B.
 submitted to the EMBL Data Library, October 1994
 A/Reference number: S48787
 A/Accession: S48788
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-428 <BUE>
 A/Cross-references: EMBL:X82275
 C/Superfamily: common tobacco sucrose transport protein

Query Match 10.4%; Score 297.5; DB 2; Length 428;
 Best Local Similarity 25.8%; Pred. No. 5.1e-15;
 Matches 114; Conservative 82; Mismatches 163; Indels 83; Gaps 18;
 Qy 4 RLVSRLRHRAQLLNNLLTGLBVLAAAGITYVPPLLLEVGVBEKFTMWLIGPVL 63
 Db 25 KLM-----KIIIVASIAAGVQGWALQSLTTPYVQLLGIPIHREFAPFIWLCGPTS 74

Qy 64 GLVCVPLGSSADHWGRGRPPFFIWNLSLGLTLLSLFLIPAGWLAGLCPDP-----R 118
 Db 75 GMTVPVGVYSDNCSSRFGRPPFFIWAAGAAVTIAVFLIGFAADL-GHASGDPLGKGSK 133
 Qy 119 PLELALLILGVLLDFCGVCFPTPLLEALLSDLFRDPD-HCRQAYSVYAFMISLGGCLGY- 176
 Db 134 PRAIAVVGFWILDVANNMLOGPCRALLDLSSGSGKWRATANAFFSFKAIGNILGYA 193
 Qy 177 -----LLPAIDWD-TSALAPYIGTQEBCLF-GLTLTILFTCVATLL-----VABEALIG 224
 Db 194 AGSYSLFKVFPFSKTKACDMYTRANKSCFFFAIFLLLSLTTLTALTVRENELEPEKELE 253
 Qy 225 PTPPAGLSAPLSPHCCPCRLARLAFNLGALLPRHLQCCRMPTLRRLFVAELCSNWA 284
 Db 254 IDEKLSGAG-----KSKVPF-FGEIFGALKDL-----PRPMWILLVTCNLNIA 296
 Qy 285 LMTFTLFTYDFVGEGLYQGVPRABPGTEARRHYDEGVMSGLGLFLOCAISLVFLVMDR 344
 Db 297 WFPFFLYDTDMWAKEVEFG-----QVGD--AKLYDLGVRAAGLGLLLQSVVLGFMFLGVBF 350
 Qy 345 LVQRFCTRAVYLASVAAPVAAGATCLSHSAVV-----TASALITGFT----- 388
 Db 351 LKKIG-GAKRLMGILNVL---AICLAWTILVTQAKSRGHDAGTLMGTPGVKIGA 406
 Qy 389 ---FSALQI-----LPYTLA 400
 Db 407 LLLFAALGIPLAVTFISIPFALA 428

RESULT 12
 S38657
 sucrose transport protein ptpl - common plantain
 N/Alternate names: sucrose transporter ptpl
 C/Species: Plantago major (common plantain)
 C/Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 26-May-2000
 C/Accession: S38657
 R/Gahrtz, M.; Stolz, J.; Sauer, N.
 submitted to the EMBL Data Library, November 1993

A/Reference number: S38657
 A/Accession: S38657
 A/Molecule type: mRNA
 A/Residues: 1-510 <GAH>
 A/Cross-references: EMBL:X75764; NID:g415987; PIDN:CAA53390.1; PID:g415988
 C/Genetics:
 A/Gene: ptpl
 C/Superfamily: common tobacco sucrose transport protein
 C/Keywords: sugar transport

Query Match 10.3%; Score 294; DB 2; Length 510;
 Best Local Similarity 24.7%; Pred. No. 1.1e-14;
 Matches 126; Conservative 77; Mismatches 199; Indels 108; Gaps 19;
 Qy 18 LLIYNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLIGPVLGLVCVPLGASDH 77
 Db 29 IFLVAAAGVQGWALQSLTTPYVQLLGIPIHKNWASYIWLCPISGMIVQPVVGVFSDN 88
 Qy 78 WRGRYGRRRPFIWALSGLIL-LSLFLIPRAGWAGL-----LCPDPRPLELALLILGV 129
 Db 89 CTSRFGRRRPF-I-AAGAGLVGVAVLI---GFAADLGHAGGDSLGDGLKPRAGVVFVG 144
 Qy 130 GLDFCGVCFPTPLLEALLSDLF-RDPDHCRQAYSVYAFMISLGGCLGYLLPAID----- 182
 Db 145 WILDVANNMLOGPCRALLDLSSGNTKMANANSFFSFVAVGNVLGYAAGSYRMYKVF 204
 Qy 183 --WTSALAPYLGTOBEC-LFGLTLTILFTCVATLL-----VAEBAALGPTPEPAGLSA 234
 Db 205 PFSKTKACDIYCANLKSCFIISITLITLITLALSIVREKRVHAEQV---TAAKGFKI 261
 Qy 235 PLSLPHCCPCRLARLAFNLGALLPRHLQCCRMPTLRRLFVAELCSNWMALMTFTLYTD 294
 Db 262 P-----VPELFGALKDLPRWVLLLTALNWIAMFGLFLFDID 301
 Qy 295 FVGBGLYQGVPRABPGTEARRHYDEGVMSGLGLFLOCAISLVFLVMDRLVQRFG--- 350
 Db 302 WNGREV-GETQHQKAPELAVIYKNGVSGALGMLNSIVLGFASLGQVMAHALGVKVR 360
 Qy 351 -----TRAVYLA-SVAAPFAAGATCLSHSAVVVTAALTG--FTFSALQI----- 394
 Db 361 LWGVNFIALICLMTITIVTKVASHRPYSGVQLQTPSSVKIGALVWFSALGIPLAITF 420
 Qy 395 -LPYTLASLTHREKQVFLPKYRGDTGGSSEDSI-----MTSFLPQPKG--- 438
 Db 421 SVFPALASIV-----STTTGSGGGLSLGVNLNLAIVTPQMIVSVASGFWMDMFG 468
 Qy 439 ---APPNGHVAGAGSG-----LLPPPPA 459
 Db 469 GGNLPFAVVGAVAAASGIFAFMTLPSPPA 498

RESULT 13
 S51114
 sucrose-proton symporter - beet
 C/Species: Beta vulgaris (beet)
 C/Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 26-May-2000
 C/Accession: S51114
 R/Westram, A.; Eckhardt, U.; Frommer, W.B.; Riesmeier, J.W.
 submitted to the EMBL Data Library, January 1995
 A/Description: Sequence of a sugar beet sucrose transporter cDNA.
 A/Reference number: S51114
 A/Accession: S51114
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-523 <WES>
 A/Cross-references: EMBL:X83850; NID:g633171; PID:g633172
 C/Superfamily: common tobacco sucrose transport protein

Query Match 10.3%; Score 294; DB 2; Length 523;
 Best Local Similarity 24.3%; Pred. No. 1.2e-14;
 Matches 125; Conservative 78; Mismatches 199; Indels 112; Gaps 20;
 Qy 17 QLLIYNLITFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLIGPVLGLVCVPLGASDH 76

Db 38 KLAIVASIAAGVQFGWALQSLSTPYVOLLGIPHTWAPYIWLCPISGMIVQPTVGYSD 97
QY 77 HWRGRYRRRPFPTWALSGLTLLSLFLIPRAGWLA--GLLCPD-----PRPLELALLILGW 129
Db 98 RCTSKFGRSPFI---AVGATLVGFVSLIGFPAADIGHATGDPNGNVKPRALIAVVGCF 154
QY 130 GLIDFCGQVCTFPLEALLSLDFRDP--HCRQAVSYVAFMISLGGCLGY-----LLP 179
Db 155 WILDVANNTLQGPCRALLADMAAGSQAKTRYANAFSFFMALGNIGGYAGSYGRLYTVFP 214
QY 180 AIDWDSALAPYLGTOBECIFGLTLTFLTCVATLVAEEAALGTEPAEGLSAPSLP 239
Db 215 FT--HTKACTYCANLKSCFFISITLIVLTIALSVRER-----PFTDEQEBENLNK 268
QY 240 HCCPCRLARLNLGALLPRLHOLCCMRPTRLRLFAELCSMWALMTTFLFYDFVREG 299
Db 269 NTGCG-ARLPF--FGQLFGALKDL----PKPMLILLVLTCLNWIAMPFLLPDDTMGKE 321
QY 300 LYQGVPAEPTGTEARRHYDEGRMGSGLGLFLQCAISLVSLVMDRLVQRF--TRAVYLA 357
Db 322 VYGGT-----VOEGKAYDMGVHAGALGLMINSVLGIMSLGIEKRLARLVGGVKRLWGI 375
QY 358 SVAAPFVAAGATCLSHSVAVVTASA-----ALTG 387
Db 376 NLIL-----AVCLAMTI-LVTXSAEYRATHVPGALGPPLPPGVKGCALAIFAVLGI 428
QY 388 TFSALQILPTLASLVHREK-----QVFLPKYRGDTGGASSEDLSMT 429
Db 429 PLAITFISIPFALASIFSASSGSGQLSLGVNLIAIVVQPMFVSVTGPW-----DAL-- 480
QY 430 SFLPGPKGAPFPNGHVAGGSG-----LLPPPP 458
Db 481 -FGGNLPA--FVGVAVATASAILSTLLPPPP 511

RESULT 14

F96741
Probable sucrose transport protein F17M19.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96741
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96741
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005173; NID:G6978914; PIDN:AAF34306.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292.5; DB 2; Length 512;
Best Local Similarity 23.8%; Pred. No. 1.5e-14;
Matches 120; Conservative 87; Mismatches 211; Indels 87; Gaps 18;

QY 17 QLLVNLITFGLVCLAAAGITYVPPLLLEVGVKEKMTVMVLGIPVLGVCVPLIGSASD 76
Db 33 KLAIVASIAAGVQFGWALQSLSTPYIQLLGIPIHKWSSYMWLCPISGMIVQPTVGYSD 92
QY 77 HWRGRYRRRPFPTWALSGLTLLSLFLIPRAGWLA-----LCPDPRPLELALLILGW 129

Db 93 RCEGRFRRRPFIAAGVALVAVSVFLI---GFAADMGHSGFDKLENKVRTRAILIIFLTGF 149
QY 130 GLIDFCGQVCTFPLEALLSLDFR--DPOHCRQAVSYVAFMISLGGCLGY-----LL 178
Db 150 WFLDVANNTLQGPCRAFLADLAAGDAKTRVANACFSFEMAVGNVGLVGAAGSYTNLHKMF 209
QY 179 PAIDWDTGALAPYLGTOBECIFGLTLTFLTCVATLVAEEAALGTEPAEGLSAPSL 238
Db 210 PFT--MTKACDIYCANLKTCCFLSITLILLIVTFSLSWYVKDK---QWSPQGDKEEKT 263
QY 239 PHCCPCRLARLNLGALLPRLHOLCCMRPTRLRLFAELCSMWALMTTFLFYDFVGE 298
Db 264 -----SLFF--FGEIFGAVR---HMKRPMWLLIIVTINNIAFPFLLIYDWMGR 309
QY 299 GLYQGVPAEPTGTEARRHYDEGRMGSGLGLFLQCAISLVSLVMDRLVQRFTRAVYLA 358
Db 310 EYVGG--NSDGDERSKLYDQVQAGALGLMFNSGILLGFLSVLGSVESIGKMG-CAKRLWG 366
QY 359 VAAFPVAAGATCLSHSVAVVT-----SNAALGTFTSALQIL----- 395
Db 367 CVNFIATG--LAWTVLVTXSAEHHREIAGPLAGSPSGIKAGVPSLFTVIGIPLAITYS 423
QY 396 -PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSLFPGPKP---GAPFPN---GHVG 447
Db 424 IPFALASIFSTNSGAGQGLSLGVNLIAICIPQMIVSFSSGFLDAQFGGNNLPFSFVGAIA 483
QY 448 AGGSGLL-----PPP--PALCGA 463
Db 484 AAVSGVLTALTVPSPPPDAPAMSGA 508

RESULT 15

T02982
Probable sucrose transport protein - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T02982
R:Hitose, T.; Imai, N.; Scofield, G.N.; Furbank, R.T.; Onaugh, R.
Plant Cell Physiol. 38, 1389-1396, 1997
A:Title: cDNA cloning and tissue specific expression of a gene for sucrose transporter
A:Reference number: Z14809; MUID:98182940; PMID:9522469
A:Accession: T02982
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-537 <HIR>
A:Cross-references: EMBL:D87819; NID:92723470; PIDN:BAA24071.1; PID:92723471
A:Experimental source: cultivar Nipponbare, leaf
C:Genetics:
A:Gene: SUT1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292; DB 2; Length 537;
Best Local Similarity 26.7%; Pred. No. 1.7e-14;
Matches 115; Conservative 72; Mismatches 167; Indels 76; Gaps 19;

QY 17 QLLVNLITFGLVCLAAAGITYVPPLLLEVGVKEKMTVMVLGIPVLGVCVPLIGSASD 76
Db 51 RLILSGMVAGGVQVQWALQSLTLPYVOTLGLSHALTSMWMLCGPIAGMVVQPCVGLYSD 110
QY 77 HWRGRYRRRPFITWALSGLTLLSLFLI---PRAGWLAGLLCPD-----PRPLELALLIL 127
Db 111 RCTSKWRRRPFYITGCVLICVAVVIGFSADIAGMDTKEDCSVHGSRWHAIVVYL 170
QY 128 GVGLDFCGQVCTFPLEALLSLDF--RDPDHCRAQSYVAFMISLGGCLGYLLPAID--- 182
Db 171 GFWLLDFSNVTVQPARALMADLSGRHGP--TANSTFCSSMWAVGNILGYSSGSTNNWH 227
QY 183 -WDSALAPYLGTOBEC-----LFG--LLTLIFIT-CVAATLLVAEE-----BALGT- 226
Db 228 KW-----FPFLKTRACCEACANLGAFLVAVIFLSLCLVITLIFAKEVPFKGNAL-PTK 281
QY 227 --EPAEGLSAPSLSPHCCPCRLARLNLGALLPRLHOLCCMRPTRLRLFAELCSWMA 284

Db 282 SNEPAEPEGTG-----PLAVLKGFRLFTGMPVS-----LIVTGL-TWLS 320
QY 285 LMTFTLLFYTDVGEGLVQGVPR-AEPGTEARRHYDEGVRMGSLGLEFQCAISLVFSLVMD 343
Db 321 WFPFLLYDMDWGREIYHGDPKGTDQIEA---FNOGVRAAGFGLLNSIVLGFSSFLIE 377
QY 344 RLVORFGTRAVYLASVAAPFVAAGATCL-----SHSVA--VVTASAALTGFTFSALQ 393
Db 378 PMCRKVGPRVWVVTNSFLVCIAAATALISFWSLKDHFHGTVOKAITADKSIKAVCLVLFA 437
QY 394 ILPYTLASLY 403
Db 438 FLGVPLAVLY 447

Search completed: August 25, 2004, 02:28:38
Job time : 51.1993 secs

This Page Blank usdo

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 01:58:37 ; Search time 29.5196 Seconds
(without alignments)
975.448 Million cell updates/sec

Title: US-09-593-793a-113
Perfect score: 2861
Sequence: 1 MVQLWVSRLLRRKQALL.....AIYFATQVVFDSKLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	489.5	17.1	530	1 MATP MOUSE	P58355 mus musculus
2	488.5	17.1	530	1 MATP HUMAN	Q9umx9 homo sapien
3	451	15.8	751	1 PAST MOUSE	Q8bi47 mus musculus
4	446	15.6	751	1 PAST RAT	Q8k4s3 rattus norv
5	324.5	11.3	525	1 SUT EPOL	Q03411 spinacia ol
6	267	9.3	509	1 PAST HUMAN	Q9y2w3 homo sapien
7	235.5	8.2	553	1 SUT1 SCHPO	O14091 schizosacch
8	142	5.0	544	1 YD74 SYNY3	P74168 synchocyst
9	136	4.8	541	1 GT10 HUMAN	O95528 homo sapien
10	125.5	4.4	399	1 TCRI ECOLI	P02982 escherichia
11	121.5	4.2	503	1 PURE STRLP	P42670 streptomyc
12	119.5	4.2	491	1 AMPG ECOLI	P36670 escherichia
13	117	4.1	440	1 YHJE ECOLI	P37643 escherichia
14	117	4.1	495	1 SMVA SALTY	P37594 salmonella
15	111.5	3.9	640	1 Y051 MYCTU	O10880 mycobacteri
16	111	3.9	473	1 PHDK NOCSK	Q24723 nocardioid
17	110.5	3.9	680	1 CALA HUMAN	Q03692 homo sapien
18	110	3.8	368	1 GALT HUMAN	O60755 homo sapien
19	110	3.8	606	1 NQCC THETH	Q50186 mycobacteri
20	109.5	3.8	654	1 SPH2 HUMAN	Q5nra0 homo sapien
21	109	3.8	477	1 YFUM RHOCA	P26176 rhodobacter
22	108.5	3.8	702	1 OPGX XANCP	Q8pdd7 xanthomonas
23	107	3.7	419	1 CMLA PSEAE	P32482 pseudomonas
24	105	3.7	448	1 VJ94 YEAST	P47159 saccharomyc
25	104	3.6	476	1 MELB SALTY	P30878 salmonella
26	103	3.6	465	1 FTSW MYCLE	Q50186 mycobacteri
27	102.5	3.6	461	1 PUCC RHOCA	P23462 rhodobacter
28	102	3.6	438	1 SHIA ECOLI	P67350 escherichia
29	102	3.6	702	1 HPPA RHORU	O68460 rhodospirill
30	102	3.6	794	1 HPPA STRCO	Q9x913 streptomyc
31	102	3.6	801	1 HPPA STRAW	Q826j8 streptomyc
32	101.5	3.5	428	1 YXIO BACSU	P42306 bacillus su
33	101.5	3.5	510	1 NANT_YERPE	Q8zch3 yersinia pe

ALIGNMENTS

RESULT 1

ID	MATP MOUSE	STANDARD;	PRT;	530 AA.
AC	P58355;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Membrane-associated transporter protein (AIM-1 protein) (Melanoma antigen AIM1) (Underwhite protein).			
GN	MATP OR AIM1 OR UW.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ddv; TISSUE=Eye, Kidney, and Uterus;			
RX	MEDLINE=21372467; PubMed=11479596;			
RA	Fukamachi S., Shimada A., Shima A.;			
RT	"Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka.";			
RL	Nat. Genet. 28:381-385(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANTS UW-DBR ASN-153 AND PRO-435.			
RX	MEDLINE=21473748; PubMed=11574907;			
RA	Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davissson M.T., King R.A., Brilliant M.H.;			
RT	"Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4.";			
RL	Am. J. Hum. Genet. 69:981-988(2001).			
CC	-!- FUNCTION: Melanocyte differentiation antigen. May transport substances required for melanin biosynthesis (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By similarity).			
CC	-!- TISSUE SPECIFICITY: Melanocytes, eyes, kidney and uterus.			
CC	-!- DISEASE: Defects in MATP are the cause of the UW-dbr phenotype that results in loss of nearly all pigmentation in the homozygous state.			
CC	-!- SIMILARITY: Belongs to the glycoside-pentoside-hexuronide (GPH) cation symporter transporter (TC 2.A.2) family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF360357; AA651713.1; -			
DR	MGI; MGI:2153040; Matp.			
KW	Melanin biosynthesis; Transmembrane; Antigen; Glycoprotein; Vision; Disease mutation; Albinism.			
FT	DOMAIN 1 45 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 46 66 1 (POTENTIAL).			
FT	DOMAIN 67 68 EXTRACELLULAR (POTENTIAL).			

34	101	3.5	437	1	BRAZ_PSEAE	P25185 pseudomonas
35	101	3.5	471	1	MELB_ENTAE	O07366 enterobacte
36	101	3.5	471	1	MELB_KLEPN	Q02581 klebsiella
37	100	3.5	404	1	YAXM_RHISN	P25705 rhizobium s
38	100	3.5	473	1	YIHO_SALTY	Q917r4 salmonella
39	99	3.5	379	1	CYB_SORHY	O79453 sorex hoyi
40	99	3.5	379	1	CYB_SORJA	Q8se00 sorex jacks
41	99	3.5	386	1	CV03_HUMAN	Q9y3p4 homo sapien
42	99	3.5	481	1	LNRA_STRLN	P46104 streptomyc
43	98.5	3.4	613	1	NUOL_ECOLI	P33607 escherichia
44	98.5	3.4	685	1	PHUB_SALTY	O87656 salmonella
45	98	3.4	379	1	CYB_SORCI	O79451 sorex ciner

```

FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 105 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 106 126 3 (POTENTIAL).
FT DOMAIN 127 138 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 6 (POTENTIAL).
FT DOMAIN 238 318 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 319 339 7 (POTENTIAL).
FT DOMAIN 340 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 8 (POTENTIAL).
FT DOMAIN 388 398 9 (POTENTIAL).
FT TRANSMEM 399 419 10 (POTENTIAL).
FT DOMAIN 420 425 10 (POTENTIAL).
FT TRANSMEM 426 446 11 (POTENTIAL).
FT DOMAIN 447 477 11 (POTENTIAL).
FT TRANSMEM 478 498 12 (POTENTIAL).
FT DOMAIN 499 504 12 (POTENTIAL).
FT TRANSMEM 505 525 12 (POTENTIAL).
FT DOMAIN 526 530 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 153 153 D -> N (IN UW-DBR).
FT VARIANT 435 435 S -> P (IN UW-DBR).
SQ SEQUENCE 530 AA; 57961 MW; 74EEDAD07916D9FC CRC64;

Query Match 17.1%; Score 489.5; DB 1; Length 530;
Best Local Similarity 26.1%; Pred. No. 1.9e-28;
Matches 146; Conservative 85; Mismatches 228; Indels 101; Gaps 12;

QY 17 QLLVLLTGLVCLAAAGTYVPPILLVGVVEKEFTMTVLGTGVLGVLCVPLLGSSASD 76
DB 34 RLNVGSHVAMGREGFCYAVEAAYVTPVLLSVGLPKSLYMWLLSPILGFLLPVVGSSASD 93

QY 77 HWGRVGRRRPFTWALSGLILLSLFLIPRAGWLAGLLCPDRP---LELAILGLVLLD 133
DB 94 HCRARWRRRPYITLAIWMLGMALYNGDAVVSALVANPRKQITWISITWGVVLPFD 153

QY 134 FCQGVCTPLEALLSLDFRDPDRCQAYSVAPMISLGGCLGVLLPAPIDWDTSLAPYL 193
DB 154 FSADFDPKIKAYLFDVCSHQDK-EKGLHYHALFTGFGALGYLLCAIDWHLDLGRLLG 212

QY 194 TQECULFGLTLIFLCAATLVLVBEAAL--GPTPEP-----AEGLSAPLSPHCCPC 244
DB 213 TEFQVWFFFSALVLLICFTIHLCSIPAPLRDAATDPPSQDPPQSSLSASGMHEY- 268

QY 245 RARLAFRNLGA-----LLPRLHQLCCRMPTLRLRLFVAELCSW 282
DB 269 GSIEKVNKGADTEQPVQEWKNNKPKSQGQORTSMKSLRALVNMPSHYRCLCVSHLGN 328

QY 283 MALMTFTLFYDFVGBGLVQGVPRABEGTEARRHYDEGVRMGSLGLFLQCA:SLVFSVM 342
DB 329 TAFLSNMLFTDFEMGQIVYHGDYGAHNSTEFLIYERGVGVGCGWGLCINSVFSSVYFQ 388

QY 343 DRLVQRFGTAIVLASVAAPVPAAGATCLSHSVAVVTAASALTGFTFSALQILPYTLASL 402
DB 389 KAMVSYIGLGLYFMGYLLFGLGTGIFGFPNNYSTLVLCSPMGVMSSTLYTVPNPLAE 448

QY 403 YHREKQVFLPKYRGDTGGASSEDLSMTSLFPGPKGAPFPNGHVAGGSGLLFPPPALCG 462
DB 449 YHREEB-----KKGQEA-----PGGPDNQR-----DCA 477

QY 463 ASACDVSVRVVGGEPTAEARVVRGICLDLAILDSAFILSQVAPSLFMGSIYQLSQSYTA 522
DB 478 ALTCMVQL-----AQILVGG-----LGLFLNMAGSVVV 506

QY 523 YMVSAAGLGLVAIFYATQVV 542
DB 507 VVITASAVSLIGCCFVALFV 526

```

```

MATP_HUMAN
ID MATP_HUMAN STANDARD; PRT; 530 AA.
AC Q9UMX9; Q9BWT3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Membrane-associated transporter protein (AIM-1 protein) (Melanoma
DE antigen AIM1).
GN MATP OR AIM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM AIM-1A).
RC TISSUE=Melanoma;
RX MEDLINE=21115844; PubMed=1121837;
RA Harada M., Li Y.F., El-Gamil M., Rosenberg S.A., Robbins P.F.;
RT "Use of an in vitro immunoselected tumor line to identify shared
RT melanoma antigens recognized by HLA-A*0201-restricted T cells.";
RL Cancer Res. 61:1089-1094(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM AIM-1B).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Max S.I., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Farber A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP ALTERNATIVE SPLICING.
RA Ferro S.;
RL Unpublished observations (NOV-2001).
RN [4]
RP DISEASE, AND VARIANT LEU-374.
RX MEDLINE=21473749; PubMed=11574907;
RA Newton J.M., Cohen-Barak O., Haglwar N., Gardner J.M., Davison M.T.,
RA King R.A., Brilliant M.H.;
RT "Mutations in the human orthologue of the mouse underwhite gene (uw)
RT underlie a new form of oculocutaneous albinism, OCA4.";
RL Am. J. Hum. Genet. 69:981-988(2001).
CC -!- FUNCTION: Melanocyte differentiation antigen. May transport
CC substances required for melanin biosynthesis (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=AIM-1a;
CC IsoId=Q9UMX9-1; Sequence=Displayed;
CC Name=AIM-1b;
CC IsoId=Q9UMX9-2; Sequence=VSP_006297, VSP_006298, VSP_006299;
CC Name=AIM-1c;
CC IsoId=Q9UMX9-3; Sequence=VSP_006296;
CC -!- TISSUE SPECIFICITY: Expressed in most melanoma cell lines and
CC melanocytes.
CC -!- DISEASE: Defects in MATP are the cause of oculocutaneous albinism

```


RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Walested C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kikuhara T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 CC -!- FUNCTION: Mediates glucose uptake along the pH gradient (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the glycoside-pentoxide-hexuronide (GPH)
 CC cation symporter transporter (TC 2.A.2) family.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 217.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AK082651; BAC38560.1; ALT_FRAME.
 CC DR PIR; PT0633; PT0633.
 CC MGD; MGI-265323; Dnbs.
 CC Transmembrane; Transport; Sugar transport; Symport.
 CC FT TRANSMEM 93 113 POTENTIAL.
 CC FT TRANSMEM 123 143 POTENTIAL.
 CC FT TRANSMEM 155 175 POTENTIAL.
 CC FT TRANSMEM 191 211 POTENTIAL.
 CC FT TRANSMEM 233 253 POTENTIAL.
 CC FT TRANSMEM 268 288 POTENTIAL.
 CC FT TRANSMEM 536 556 POTENTIAL.
 CC FT TRANSMEM 576 596 POTENTIAL.
 CC FT TRANSMEM 606 626 POTENTIAL.
 CC FT TRANSMEM 630 650 POTENTIAL.
 CC FT TRANSMEM 688 708 POTENTIAL.
 CC FT TRANSMEM 710 730 POTENTIAL.
 CC SQ SEQUENCE 751 AA; 81420 MW; 8242480E154F6BE1 CRC64;
 Query Match 15.8%; Score 451; DB 1; Length 751;
 Best Local Similarity 21.7%; Pred.No. 1.8e-25;
 Matches 155; Conservative 80; Mismatches 212; Indels 268; Gaps 13;
 QY 17 QLLVLLNLTGELVCLAAAGTYVPPILLEVGEVEKFTWVLGIPVLGVLCVPLLGASD 76
 DB 86 ELFLNGCILGIFSYAMETAYTPVLLQNGLDQLYSLWFLSPILGFLQLLQGLGAWSD 145
 QY 77 HWRGYRRPPTWALSGLILLSFLIPRAGWAGLLCPDP-----RPLEALLILGVGLLD 133
 DB 146 RCTSRFRGRRPFFILVLAIGALLGLSLLN-GRDIGMALADTATNHWGILLTVCVGLMD 204
 QY 134 FCGQVCFTPEALLSDFRPHDCRQAYSVYAFMISLGGCLGYLLPAIDWDTSALAPYLG 193
 DB 205 FSADSANPNHAXMWDVCGPVDQR-GLNHALMAGLGGFGVVGGIHWKTSFGRALG 263
 QY 194 TQECLEGLITLIFLCVATLIVABEALGPT-EPAGELGASPL----- 237
 DB 264 GQLRVIVFTAITLSVTTLTISIPERPLRPLGKRTAMKSPSLPPLPPVILBEGAG 323
 QY 238 -----SPHCCP----- 243
 DB 324 DALPSTTATSLIYAFSSPISPPSLPKYGSFISROSLTGINEFPASSFGTNSIDSLVD 383
 QY 244 -----CRALAFRNIGALL----- 257

DB 384 CFTAGHDNYLALPSSVPRQAISVSFPRAPDGFYCOERGLERREGPLTLGSDGDLVRVGLS 443
 QY 258 ----PR----- 259
 DB 444 DTSKPRASGILKRPTQTLALPDVAGGNGPETSRRNRVTFSQVANILINGVKYSELTGSS 503
 QY 260 -----LHQGCC---RMPRTLRLEVAELCSMMALMTFTLVTPVGGSLYQGVPRAE 308
 DB 504 EQSQPJSRLRHCSTIYNMFKALRNCVNHFLGWSPEGMLLFTYTPMGVFWFGDPRAP 563
 QY 309 PGTERRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFGRVAVYLAASVAFPPVAAGA 368
 DB 564 HTSEAYQKNSGVTMGCGMCIIVAFSAFSAILEKLECLSVRTLYFIAYLAFGLGTGL 623
 QY 369 TCLSHVAVVTASALGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDLSL 428
 DB 624 ATLSRLNYLVLSLCTTYGILFTLCTLPYSLLCDYQSK-----KFAG-----SSADGTR 673
 QY 429 TSFLPGKPGAPFPNGHVGAGGSLPPPALCGASACDVSVRVVVGEPTEARVVPGRGI 488
 DB 674 -----RGM 676
 QY 489 CLDAILDSAPLLSQVAPSLFMGSIVQLSOSVTVYMWVAAGLGLVLIYFATQVVF 543
 DB 677 GVDISLLSCQVFLAQILVSLVLP-----LTSVAGSANGVMYFSSLSVVF 720
 RESULT 4
 PAST RAT STANDARD; PRT; 751 AA.
 AC Q8K4S3;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Proton-associated sugar transporter A (PAST-A).
 GN DNB5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22305182; PubMed=12417639;
 RA Shimokawa N., Okada J., Haglund K., Dikic I., Koibuchi N., Miura M.;
 RT "Past-A, a novel proton-associated sugar transporter, regulates
 RT glucose homeostasis in the brain.";
 RL J. Neurosci. 22:9160-9165(2002).
 CC -!- FUNCTION: Mediates glucose uptake along the pH gradient.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in brain.
 CC -!- SIMILARITY: Belongs to the glycoside-pentoxide-hexuronide (GPH)
 CC cation symporter transporter (TC 2.A.2) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB075229; BAB97313.1;
 DR Transmembrane; Transport; Sugar transport; Symport.
 KW TRANSMEM 93 113 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 191 211 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 268 288 POTENTIAL.
 FT TRANSMEM 536 556 POTENTIAL.
 FT TRANSMEM 576 596 POTENTIAL.
 FT TRANSMEM 606 626 POTENTIAL.


```

356 LASVAAPFVAGATCLSHSVAVVYASAAALTGTFSSALQILPYTLASLYHREKQVFLPKRY 415
376 I V N I I L -----AVCLAMTV-LVTKSAE-----HFEDSHHIM----- 405
416 GDTGGASSEDLSMTSLFPGKPGAPFPNGHVAGGSGGLPPPPALCGASACDVSVGVVVG 475
406 -----GSAVFPFPPPA--GVKGGALATFAVLG 429
476 EPTEARV-VP-----GRGICLDLAILDSAPFLSQV-----AP 506
430 IPLAIFSIFFALASIFSASSGGQGLGLVNLAIWVPMFVSVTSGPMDAMFGGNNLP 489
507 SLFMGSIVQLSQSVTAY 523
490 AFVGVGAATAASAVLSF 506

RESULT 6
PAST_HUMAN STANDARD; PRT; 509 AA.
ID PAST_HUMAN
AC Q9Y2W3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Proton-associated sugar transporter A (PAST-A) (Deleted in
DE neuroblastoma 5 protein) (DNb-5) (Fragment).
GN DNb5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20195630; PubMed=10729226;
RA Anler L.C., Bauer A., Corvi R., Dohlmann S., Praml C., Cavenee W.K.,
RA Schwab M., Hampton G.M.;
RT "Identification and characterization of novel genes located at the
RT t(1;15)(p36.2;q24) translocation breakpoint in the neuroblastoma cell
RT line NGP.";
RL Genomics 64:195-202(2000).
CC -!- FUNCTION: Mediates Glucose uptake along the pH gradient (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in adult heart, brain, muscle and
CC kidney, with very strong expression in brain. Also expressed in
CC fetal brain, kidney and lung.
CC -!- SIMILARITY: Belongs to the glycoside-pentoxide-hexuronide (GPH)
CC cation symporter transporter (TC 2.A.2) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF118274; RAD27583.1; ALT_INIT.
CC Genew; HGNC:17939; DNb5.
CC MTM; 605763;
KW Transmembrane; Transport; Sugar transport; Symport.
FT NON_TER 1 1
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
FT TRANSMEM 469 489 POTENTIAL.
SQ SEQUENCE 509 AA; 54637 MW; EEE766819D267C4C CRC64;

```

Query Match

9.3%; Score 267; DB 1; Length 509;

```

Best Local Similarity 25.3%; Pred. No. 3.2e-12;
Matches 72; Conservative 42; Mismatches 93; Indels 78; Gaps 4;
259 RLHQLCCRMPTLRRLRFAELCSWMLMALTFTLFTVDFVGGELYQGVPRAPGTEARRHYD 318
272 RLCTICNMPKALRTLCVNHFLGWLSEFGLMLFTVDFMGEVVFQGDPKAPHTSEAYQKYN 331
319 EGVFMGSLGFLQCAISLVSLVMDRLVQRFGRTRAVYLASVAAPFVAGATCLSHSVAVV 378
332 SGVTMGWGMCIYAFSAFYSAILEKLEELSVRTLYFIAYLAFGLGTGLATLSRLNLYV 391
379 TASAALGTFFSALQILPYTLASLYHREKQVFLPKRGDGTGGASBDSLSMTSLFPGKPG 438
392 LSLCITYGILFSTLCITLPYSLLCDYQSK-----KPAQ-----SSADGTR----- 431
439 APFNGHVAGGAGGLPPPPALCGASACDVSVGVVVGTEARVVPFGRCGLDLAILDSA 498
442 -----RGMGVDSILSCQ 444
499 FLLSOVAPSLFMGSIVQLSQSVTAYMVVSAAGLGLVAIYFATQVVF 543
445 YFLAQILVSLVGP-----LTSVGSANGVMYFSSIVSF 478

RESULT 7
SUT1_SCHPO STANDARD; PRT; 553 AA.
ID SUT1_SCHPO
AC O14031; Q9U060;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE General alpha-glucoside permease.
GN SUT1 OR SPAC2f3.08.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Weltjens I., Vanstreels E., Aert R., Robben J., Grymonprez B.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Lehrach H., Reinhardt R., Pohl T.M.,
RA Borzym K., Langer I., Beck A., Lebrach H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Dominguez A., Revuelta J.L., Moreno S., Sanchez M., del Rey F., Benito J.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrall B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 35-166 FROM N.A.
RC STRAIN=968 H90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,

```



```

Db      149 SYLLDVA NVNVMASTRSLIVDSVRS-DQOHEANSWAGRMIGVNLGYLGLYLIRIF 207
QY      187 ALAPYLGTGEELFLGLTLTFTCAATLLVAEEAALGPTPEAGLSAPSLSPHCCPRA 246
Db      208 SPLNFTQLQVFCVLASISLV-LTVTTITFVSR-RFPVHEKSVAG----- 253
QY      247 RLAFRNGLALLPRHLQCCRMPTIRRLFAVELCSWALMTFLFLTFDFVQEGLYQGVPR 306
Db      254 -----EIFEFTTMRQSIATLPTFLKRICFVQFFAYFGNFFFLFYITTYGVILYLRHAPK 308
QY      307 ABPGTEARRHYDGVNMGSLGLFLQCAISLVFSLVMDRLVQ 347
Db      309 GH-----BEDWDWATQGSFALLLFAIISLAANTALPLLE 344

RESULT 8
ID_YD74_SYN3
ID_YD74_SYN3 STANDARD; PRT; 544 AA.
AC P74168;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical symporter s11374.
GN SLL1374.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC CC -1- SIMILARITY: BELONGS TO THE SODIUM:GALACTOSIDE SYMPORTER FAMILY
CC (SGP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90912; BAA18257.1; -.
DR PIR; S75696; S75696.
DR InterPro; IPR001927; Na/Gal_symport.
DR TIGRfam; TIGR00782; Gph, 1.
DR PROSITE; PS00872; NA_GALACTOSIDE_SYMPT; FALSE_NEG.
KW Hypothetical protein; Transport; Transmembrane; Symport;
KW Complete proteome.
FT FT TRANSMEM 31 51 POTENTIAL.
FT FT TRANSMEM 52 72 POTENTIAL.
FT FT TRANSMEM 84 104 POTENTIAL.
FT FT TRANSMEM 116 136 POTENTIAL.
FT FT TRANSMEM 162 182 POTENTIAL.
FT FT TRANSMEM 191 211 POTENTIAL.
FT FT TRANSMEM 230 250 POTENTIAL.
FT FT TRANSMEM 257 277 POTENTIAL.
FT FT TRANSMEM 318 338 POTENTIAL.
FT FT TRANSMEM 356 376 POTENTIAL.
FT FT TRANSMEM 383 403 POTENTIAL.
FT FT TRANSMEM 407 427 POTENTIAL.
FT FT TRANSMEM 450 470 POTENTIAL.
FT FT TRANSMEM 501 521 POTENTIAL.

```

```
SQ SEQUENCE 544 AA; 60064 MW; C91D0EDFF32277EE CRC64;
Query Match 5.0%; Score 142; DB 1; Length 544;
Best Local Similarity 22.3%; Pred. No. 0.0045;
Matches 103; Conservative 60; Mismatches 179; Indels 126; Gaps 20;
QY 16 AQLLLVLLTFLGVEVLAAG-----ITYVPLLELV-GVBEKFTWVLGIGPYL 63
DB 2 SQSLSAEKLHFTTKLAYGAGDFGPAITANILVFYLLFLFTDVAGIPAAAGVLMIGKIF 61
QY 64 GLVCPVLLGSASDHWGRYCRPPFTWALSGLILLSFLIPRAGWLAGLLCPDPRLELA 123
DB 62 DAINDPIIGLSLDRSTRSGRRLPFW-----LGMIFPFLFYTAQWLPHFSDRDTNQG 117
QY 124 LLILGVLLDFCGQVCF-----PLEALLDFRDPDCHQOAVSVAFMISLGGCGLYLP 179
DB 118 LFIYVAIA-MAFNLCVTVNLYPTALTPELTQYNE-RTRLNSPFAFISIGSILSLIL 175
QY 180 AIDWTSALAPYLGTEBECLFGLL-TLIFLTCVATLLVAEALGFTPEAGLSAPSLS 238
DB 176 YI-----LIAAGLPDRPQQFGELGVMISVLISALLWSALRIQKKEP-----ILSPSL- 226
QY 239 PHCCPCRARLA-----FRNLGAL----- 256
DB 227 -----RRRLAPLLMAAGITLILATAKSFNLLGSGFDYISFFLLILLGLINGGFGFTLR 280
QY 257 -----LPRHLQCCRMPTLR-LFV-AELCSWALMTFTLF 291
DB 281 DSAVEEHLQLENSPSPGVNTENPLKQL-----KIAFSNRAFLFVIGYLCSLAVQLTASI 338
QY 292 YTFVCGELVQGVPRPEGTGARRHYDEGRVMSGLFLQCAISLVFSVMDRLVQRFCT 351
DB 339 LTVFVVS--WNGNEQOSGT-----IALVQ-GTALVMLFVWQALAQFLDK 381
QY 352 RAVY-LASVAAPFAAGATCLS-HSVAVVTASAAALTGFTFSALQLDLYTL 399
DB 382 KVIYFLGVMVWMAEAGLMVQGVQVALLVLAIFAGVGVSVAYLIPWSM 431
RESULT 9
GT10 HUMAN
ID AC Q95528; Q9H416; STANDARD; PRT; 541 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 10
DE (Glucose transporter type 10).
GN SLC2A10 OR GLUT10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21145593; PubMed=11247674;
RA McVie-Wylie A.J., Lamson D.R., Chen Y.T.;
RT "Molecular cloning of a novel member of the GLUT family of
RT transporters, SLC2A10 (GLUT10), localized on chromosome 20q13.1: a
RT candidate gene for NIDDM susceptibility.";
RL Genomics 72:113-117(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Stavrides G.S., Hashim Y., Huckle E.J., Deloukas P.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Fossey S.C., Mihic S.J., Craddock A.L., Mychaleckyj J.C., Dawson P.A.,
RA Bowden D.W.;
RT "GLUT10: a novel glucose transporter in the type 2 diabetes linked
RT region of chromosome 20q12-13.1.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=2163749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaeslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.D., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: Facilitative glucose transporter (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed; highest levels in liver and
CC pancreas.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTERS SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL AF321240; AAK26294.1; -
EMBL AF321240; CAB69822.2; -
EMBL AF248053; AAK31911.1; -
EMBL AL031055; CAA1926.2; -
Genew; HGNC:13444; SLC2A10.
MIW; 606145; -
GO; GO:0016021; C: integral to membrane; NAS.
GO; GO:0005351; F: sugar porter activity; NAS.
GO; GO:0015758; P: glucose transport; NAS.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub transporter.
InterPro; IPR005829; Sug transporter.
InterPro; IPR003663; Sugar transpt.
Pfam; PF00083; sugar tr; 1
PRINTS; PR00171; SUGRTNSPOT.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Transport; Sugar transport; Transmembrane; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 1 (POTENTIAL).
FT DOMAIN 37 48 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 49 69 2 (POTENTIAL).
FT DOMAIN 70 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 98 3 (POTENTIAL).
FT DOMAIN 99 106 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 107 127 4 (POTENTIAL).
```

FT DOMAIN 128 134 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 135 155 5 (POTENTIAL).
FT DOMAIN 156 166 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 167 183 6 (POTENTIAL).
FT DOMAIN 188 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 284 7 (POTENTIAL).
FT DOMAIN 255 289 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 270 290 8 (POTENTIAL).
FT DOMAIN 291 298 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 299 319 9 (POTENTIAL).
FT DOMAIN 320 414 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 415 435 10 (POTENTIAL).
FT DOMAIN 436 445 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 446 466 11 (POTENTIAL).
FT DOMAIN 467 476 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 477 497 12 (POTENTIAL).
FT DOMAIN 498 541 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 541 AA; 56911 MW; 60644525FA136908 CRC64;

Query Match 4.8%; Score 136; DB 1; Length 541;
Best Local Similarity 21.9%; Pred. No. 0.012;
Matches 130; Conservative 72; Mismatches 210; Indels 182; Gaps 27;

QY 16 AQLLVNLTFTGLEVCCLAGITVPPLELVG---VEEKFTMTWLGIGVILGVCVPLIG 72
DB 12 ASVSLGGITFTGELAVISGA--LPLQLDFGLSCLEQEFVINGLLGALLASIVGGFLI 69
QY 73 SASDHWGRYGRRRPFIALSIGLLSLFLI PRAGWLAGLCPDRPLELALLIIGVLL 132
DB 70 DC-----YGRQALIGS-NLVLLAGSLTLGLAG-----SLAWVLGRAVV 108
QY 133 DFCGVCFPTLEALLSDLPDPCHCRQAYSVYAFMISLG---GCLGYLLPALDWDTS-- 186
DB 109 GFAISLSSMACCIYVSELV-GPQRGVVLSYEAGITVGILLSYALNYALAGTPGWRHM 167
QY 187 ---ALAPYLTQBECLFGLTLITFLTCVAATLLVAEEAALGTEPAEGLSAPSLPHCP 243
DB 168 FGWATAP-----AVLQSLSLFLP--ACT---DETATHKDLIPLQCGEAFKLGFG-- 212
QY 244 CRARLAFNLGALLPKHOLCCMRPTLRFLVAEIICSMWALMTFLFTDFVGEGLYQG 303
DB 213 -RPRYSFLDLFRARDNMRG---RTTVGLGLVLFQQLTGQPNVLCYA--STIFSSVGFHGG 266
QY 304 VPRAEPTGTEARRHYDEGVMSGLGLFQCAISLVFSLVMDRAVOREGTRAVYLASVA--- 360
DB 267 -----SSAVLASVGL---GAVKVAATLTANGLVDRAGREALLAGCALWA 308
QY 361 -----APFVAGATCLSHSVAVVTASALTG---FTFSAIQILPYTLASLYHR 405
DB 309 LSVSGIGLVSAFVMSGPSCL--AVENATGQTGLFGDSGLLQDSSLPIPRPRT-----N 360
QY 406 EKQVFLPKYRGDTGGASSEDLSMTSLFPGKPCAPFENGHVAGSGSLPPPPALCGASA 465
DB 361 EDQ-----REPILSTAKKTRPHRSGDPSAPPRIALSSALPGPP----- 399
QY 466 CDVSVRVVVGTEARVVPGRG-----ICLDLAILDLSAF-----LLSQVAP-- 506
DB 400 -----LPARGHALLRWLTALLCLMVVSFAFSFGFGPVTWLVLSIYPVE 442
QY 507 -----SLFMG-----SIVQLSQSVTAMWSAA-GLGIVAIY 536
DB 443 IGRGAFACNSFNWAANLFTSLSFLDLTIGTIGLSWTFLYGLTAVLGLGFIYLF 496

RESULT 10
TCR1-ECOLI STANDARD; PRT; 399 AA.
AC P02982;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tetracycline resistance protein, class A (TEF1A(A)).

GN TETA.
OS Escherichia coli.
OG Plasmid RP1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RC TRANSPONSON=Tnl721;
RX MEDLINE=92192465; PubMed=1312499;
RA Allmeier H., Cresnar B., Greck M., Schmitt R.,
RT "Complete nucleotide sequence of Tnl721: gene organization and a
RL novel gene product with features of a chemotaxis protein.";
[2]
RP Gene 111:11-20(1992).
RC SEQUENCE OF 85-399 FROM N.A.
RC STRAIN=DH1;
RA Fletcher J.N., Hart C.A., Batt R.M., Saunders J.R.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC PLASMID=RP1; TRANSPONSON=Tnl721;
RX MEDLINE=93299270; PubMed=6310527;
RA Waters S.H., Rogowsky P., Grinstead J., Altenbuchner J., Schmitt R.;
RT "The tetracycline resistance determinants of RP1 and Tnl721:
RL nucleotide sequence analysis.";
[4]
RP Nucleic Acids Res. 11:6089-6105(1983).
TOPOLOGY.
MEDLINE=92388137; PubMed=1517220;
RA Allard J.D., Bertrand K.P.;
RT "Membrane topology of the pBR322 tetracycline resistance protein.
RT TetA-PhoA gene fusions and implications for the mechanism of TetA
membrane insertion.";
RL J. Biol. Chem. 267:17809-17819(1992).
CC -!- FUNCTION: Resistance to tetracycline by an active tetracycline
CC efflux. This is an energy-dependent process that decreases the
CC accumulation of the antibiotic in whole cells. This protein
CC functions as a metal-tetracycline/H(+) antiporter.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X61367; CAA43643.1; -
CC EMBL; L29404; AAA83545.1; -
CC EMBL; X00006; CAA24909.1; -
CC PIR; A03509; YTECR1.
CC PIR; JQ1479; JQ1479.
CC InterPro; IPR004734; Drug resist.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub transporter.
CC InterPro; IPR005829; Sug transporter.
CC InterPro; IPR001958; TCR_TetA.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR01035; TCR_TetA.
CC TIGRFAMs; TIGR00880; 2.A_01_02; 1.
CC PROSITE; PS00850; MFS; 1.
CC PROSITE; PS00216; SUGAR TRANSPORT 1; UNKNOWN 1.
CC Antibioc resistance; Transmembrane; Inner Membrane; Transport;
KW Antiport; Ion transport; Hydrogen ion transport; Transposable element;
KW Plasmid.
FT DOMAIN 1 7 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 8 27 PROBABLE.
FT DOMAIN 28 45 PERIPLASMIC (PROBABLE).
FT TRANSMEM 46 66 PROBABLE.

```

FT DOMAIN 67 79 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 80 100 PROBABLE.
FT DOMAIN 101 103 PERIPLASMIC (PROBABLE).
FT TRANSMEM 104 124 PROBABLE.
FT DOMAIN 125 138 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 139 159 PROBABLE.
FT DOMAIN 160 160 PERIPLASMIC (PROBABLE).
FT TRANSMEM 161 181 PROBABLE.
FT DOMAIN 182 210 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 211 231 PROBABLE.
FT DOMAIN 232 246 PERIPLASMIC (PROBABLE).
FT TRANSMEM 247 267 PROBABLE.
FT DOMAIN 268 277 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 278 298 PROBABLE.
FT DOMAIN 299 299 PERIPLASMIC (PROBABLE).
FT TRANSMEM 300 320 PROBABLE.
FT DOMAIN 321 339 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 340 360 PROBABLE.
FT DOMAIN 361 364 PERIPLASMIC (PROBABLE).
FT TRANSMEM 365 385 PROBABLE.
FT DOMAIN 386 399 CYTOPLASMIC (PROBABLE).
FT CONFLICT 5 5 R -> I (IN REF. 3).
FT CONFLICT 55 55 M -> V (IN REF. 3).
FT CONFLICT 75 75 V -> I (IN REF. 3).
FT CONFLICT 84 84 A -> T (IN REF. 3).
FT CONFLICT 201 203 ASP -> SFV (IN REF. 3).
SQ SEQUENCE 399 AA; 42240 MW; 29842756BS478374 CRC64;

Query Match
Best Local Similarity 22.3%; Pred. No. 0.052;
Matches 104; Conservative 57; Mismatches 136; Indels 169; Gaps 27;

QY 19 LLYNLTGFEVCLAGIIVVPL---LLEVGVEKEFT---MWLGIVGLVGVPL 71
Db 7 LVLSTVALD---AVGILMPVLPGLRLDLVHSDNVTAHYGLIALYALMQFACAPVL 63
QY 72 GSASDHWGRGRR-----RPFIALSLGILSLFLIPRAGMLAGL 112
Db 64 GALS D---RFGRRPVLLVSLAGAAVDYAIMATAPFLWLYIGRIVA----- 106
QY 113 LCDPRLPLALILGVGLDFCGVCFTPLLEALLSDLFRDPDHCQAVSVVAFMISLGG 172
Db 107 -----GITGATGAVA---GAYIADITDGERARH-----FGFM---SA 138
QY 173 CLGYLLPAIDWDTSA LAPVLTQECLEGLLTLITFCVAATLLVAEEAALGTEPAEGL 232
Db 139 CFQFGMVA-----GPVLG-----GLM----- 154
QY 233 SAPSLSHCCPCARIAFRNLG-----ALLPRLHQCCMPRTLRLFVAELCS--W--- 282
Db 155 --GGFSPH-APPFAAALAGNLTGCFLLPESHK---GERRPLRREALNPLASPRWARG 208
QY 283 --MALMTPLTFYDFVGLGVQVRA--EPTGARRHYDE--GVRMGSLGLFLQCAI 334
Db 209 MTVVAAALMAVFFIMQLVUGQ-----VPAALWVIFGEDRFHWDATTIGISLAAPGILHSLAQ 263
QY 335 SLVFLSMDRLVQRFQTRAVYLASVA---AFFVAAGAT--CLSHSVAVVVTASAAALTGFTF 389
Db 264 AMITGPVAARLGER---RALMLGMIADGTGYILLAFATRGWMAFFIMVLLASG---GIGM 317
QY 390 SALQILPYTLASLYHREKQVFLPKRGDTGGGASDSLSLTSFLPGP 435
Db 318 PALQAM---LSRQVDEERQ-----GQLQGSAAALTSLSIV-GP 352

```

RESULT 11

```

PUB8 STRLP STANDARD; PRT; 503 AA.
AC P42670;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Puromycin resistance protein pur8.

```

```

GN PUR8.
OS Streptomyces lipmanii (Streptomyces alboniger).
CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC Streptomycineae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=132472;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12461;
RX MEDLINE=94109397; PubMed=7916693;
RA Tercero J.A., Lacalle R.A., Jimenez A.;
RT "The pur8 gene from the pur cluster of Streptomyces alboniger encodes
RT a highly hydrophobic polypeptide which confers resistance to
RT puromycin."
RL Eur. J. Biochem. 218:963-971(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVE PUROMYCIN EFFLUX ENERGIZED BY
CC A PROTON-DEPENDENT ELECTROCHEMICAL GRADIENT. IN ADDITION, IT COULD
CC BE IMPLICATED IN SECRETING N-ACETYLPUROMYCIN. THE LAST
CC INTERMEDIATE OF THE PUROMYCIN BIOSYNTHESIS PATHWAY, TO THE
CC ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL: X76855; CAA54186.1; -
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub-transporter.
DR InterPro; IPR001411; TCR-TetB.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR01036; TCRTEB.
DR TIGFAMS; TIGR00711; efflux_EmrB, 1.
DR PROSITE; PSS0850; MFS; 1.
KW Antibiotic resistance; Antibiotic biosynthesis; Transmembrane;
KW Transport.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 25 45 POTENTIAL.
FT DOMAIN 46 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 POTENTIAL.
FT DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 113 POTENTIAL.
FT DOMAIN 114 122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 123 143 POTENTIAL.
FT DOMAIN 144 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 173 POTENTIAL.
FT DOMAIN 174 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 202 POTENTIAL.
FT DOMAIN 203 212 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 213 233 POTENTIAL.
FT DOMAIN 234 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 262 POTENTIAL.
FT DOMAIN 263 280 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 281 301 POTENTIAL.
FT DOMAIN 302 313 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 314 334 POTENTIAL.
FT DOMAIN 335 346 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 347 367 POTENTIAL.
FT DOMAIN 368 371 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 372 392 POTENTIAL.
FT DOMAIN 393 422 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 423 443 POTENTIAL.
FT DOMAIN 444 461 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 462 482 POTENTIAL.
FT DOMAIN 483 503 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 503 AA; 51852 MW; FA3B814DF9A3EB74 CRC64;

```

```
Query Match      4.2%; Score 121.5; DB 1; Length 503;
Best Local Similarity 23.2%; Pred. No. 0.13;
Matches 131; Conservative 61; Mismatches 175; Indels 197; Gaps 32;

QY 26 FGLLEVCIAAGITV-----PDLLEVGVEEKEFTMW-----LGIGPVLGLVCVPL 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 WGLVVLIAQLLVLDGTVVNIALPSVQVDLGSQTSROWVITAYTAFGGL-----L 79

QY 71 LGSASDHWGR-----YGRRRPFIALSLGILLFLIPRAGW-LAGLL---CPDRPRL 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 LG-----GRVADAFGRIRF-----AVGIL-----GFLASLLGGAAPDGTLLF 119

QY 123 ALLILGVGLDFCGVCFTPLEALLSDLPDPCRCQVSVYAFMISLGGCLGYLLPAID 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 ARALQGV-----FAALAPAL-ALINTLTERGERGKAFGVGVGSGGAAVGLAG--- 171

QY 183 WDTSLAPVLTQOECLFGLLT--LIFLTCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 -----GLLTYELDWRWCLVYNAPVALLIG----- 197

QY 241 CCPCRLAFLRNLCALLPRLHQLCCMPRTLRLFAVEL--CSWMLMTFTLYFTDFVGE 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 ---CR-----LLPDR-----RTGRAVRDLDPGLLGGGLVAIVAF----- 232

QY 299 GLYQGVPRAPGPEARRHYDEG--VRMSGIGLFLQCAISLVFSLVMDRLVQRFGRVAVL 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 -----AEESGWDPLVVLVLLVGLVLMVAFALVRRVQDPLP----- 271

QY 357 ASVAAPVAGATCLSHSVAVVTASALTG-----FTFSALQILPYT--LASLYHREKQVF 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 PGVVAHRVRRGS-----FLVVGFLPQIGLFLFLTYLQGLDYSFVLTVG-----AF 320

QY 411 LPKYRGDTGASSSESLMTSFLPGPKP-----GAPFNGHVAGGSGGL-----PPPPALC 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 LPLGLGIANGSS---LIAARLLPRTPLRLLIVGALL-----AAAGWALLTRLEPDT 370

QY 462 GASACDVSVVVVGTEPTEARVPGRI-CLDLAIDSAFLLSQVAP-----SLFMGSIV 514
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 -----QVYLTHLLPAQILIGLIGCGWMPAMHTA--TARVAPHEGAAAAVVSQA 419

QY 515 QLSQSQTAYMVSAAAGLVAIYEA 538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 QVGGALGVALLNTVSTGATAAYLA 443
```

RESULT 12

```
AMPG_ECOLI
ID AMPG_ECOLI STANDARD; PRT; 491 AA.
AC P36670;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE AmpG protein.
GN AMPG OR B0433 OR Z0536 OR ECS0487.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP MEDLINE=94049112; PubMed=8231804;
RX Lindquist S., Weston-Hafer K., Schmidt H., Pul C., Korfmann G.,
RA Erickson J., Sanders C., Martin H.H., Normark S.;
RT "AmpG, a signal transducer in chromosomal beta-lactamase induction.";
RL Mol. Microbiol. 9:703-715(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SN0301-1, SN0301-3, and SN0301-5;
RX MEDLINE=95291453; PubMed=7773404;
RA Schmidt H., Korfmann G., Barth H., Martin H.H.;
RT "The signal transducer encoded by ampG is essential for induction of
chromosomal AmpC beta-lactamase in Escherichia coli by beta-lactam
```

```
antibiotics and 'unspecific' inducers.";
RL Microbiology 141:1085-1092(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Arodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21152331; PubMed=11258796;
RA Hayashi T., Makino K., Ochiaishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Probably acts as a permease in the beta-lactamase
induction system and in peptidoglycan recycling.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Probable).
CC -!- SIMILARITY: TO H.INFLUENZAE HI0350 AND TO YEAST YBR220C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL, S67816; AAB28984.1; -
CC EMBL, X82158; CAA57851.1; -
CC EMBL, X82159; CAA57852.1; -
CC EMBL, X82160; CAA57853.1; -
CC EMBL, AE000149; AAC73536.1; -
CC EMBL, U82664; AAB40189.1; -
CC EMBL, AE005222; AAG54783.1; -
CC EMBL, AP002551; BAB33910.1; -
CC PIR, C85540; C85540.
CC PIR, G90689; G90689.
CC PIR, S37391; S37391.
CC EcoGene; EG12183; ampG.
CC InterPro; IPR004752; AmpG_permease.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC Pfam; PF00083; sugar_tr; 1.
CC TIGRFAMs; TIGR00901; 2A0125; 1.
CC PROSITE; PS50850; MFS; 1.
```


QY 112 LLCPPDRP-LELALLILGVGLDFGQVCFPLEALLSLDFRDPDCHRCQAYSVYAFMISL 170
Db 342 ----SGNPLVFAFLLGLSLMG----LTFGPMGALLPELF--PTEVR--YTGASFYSNV 389
QY 171 GGCGVLLPAIDWDSALAPYGTQBECLFGLLTL-IFLTCVAAILLVA 218
Db 390 ASILG-----ASVAPYIAANLQTNVGLGAVGLYLAAMAGLTLIA 428

RESULT 14
SWVA_SALTY STANDARD; PRT; 495 AA.
AC P37594;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methyl viologen resistance protein smva.
GN SWVA OR STM1574.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
ON NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SL1303;
RX MEDLINE=95011654; PubMed=7926834;
RA Hongo E., Morino M., Mita K., Machida I., Hama-Inaba H., Tsuji H.,
RA Ichimura S., Noda Y.;
RT "The methyl viologen-resistance-encoding gene smva of Salmonella
typhimurium";
RL Gene 148:173-174(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2";
RL Nature 413:852-856(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Potential).
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D26057; BAA05055.1; --
CC ENBL; AE008769; AL20492.1; --
CC SyGene; SG10384; smva.
CC InterPro; IPR001411; MFS.
CC InterPro; IPR001411; TCR.TetB.
CC PRINTS; PR01036; TCR.TETB.
CC PROSITE; PS50850; MFS; 1.
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 5 25 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 220 240 POTENTIAL.

FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 469 489 POTENTIAL.
FT CONFLICT 144 144 A -> R (IN REF. 1).
FT CONFLICT 147 147 G -> V (IN REF. 1).
FT CONFLICT 182 182 V -> D (IN REF. 1).
FT CONFLICT 198 198 A -> V (IN REF. 1).
FT CONFLICT 227 227 F -> Y (IN REF. 1).
FT CONFLICT 447 447 L -> Y (IN REF. 1).
FT CONFLICT 454 457 AILD -> GXYLT (IN REF. 1).
SQ SEQUENCE 495 AA; 52134 MW; 961F77C748CDAL164 CRC64;
Query Match 4.1%; Score 117; DB 1; Length 495;
Best Local Similarity 21.0%; Pred. No. 0.27;
Matches 92; Conservative 60; Mismatches 151; Indels 136; Gaps 22;
QY 17 QLLVNLITLGL-----EVCLAAGTYVPPLLLEVG-----EEKFMT 54
Db 72 RLLMLGGTFLGLASLAFAASHTASWLIATRVLLAIGAAIVPATL-AGIRATFCEEKHN 130
QY 55 MVLGI-----GPVLG-----LVCVPL-----LGSASDHWGRYGR- 85
Db 131 MALGVAAVSGGAAPGLIGLLEHFYGVSVFLINPVLVVMGLTARYVPEQAGRD 190
QY 86 RPF-----IWALSL-----GILLSLFLIPRAGWLGL 112
Db 191 QPLNLGHAVMLIIAILLVVSATKALGHLSLWISFTLLTGLLGLGFIQTQATSRPM 250
QY 113 LCPDPRPLEALLILGVGLDFGQVCFPLEALLSLDFRDPDCHRCQAYSVYAFMI- 168
Db 251 I--DMRLFTTHRIILSGV-VMAVTAMITLVGFELLMAQELQFV-HGLSPYAGVFMPLVMV 306
QY 169 -----SLGGCL-----GYLLPAIDWDSALAPYGTQBECLFGLLTLIFLTC 210
Db 307 ASGSGPIAGVLVSRGLRLVATGSMALSALSYFLGAMTDP-STQQQWAGLMALLGFS- 364
QY 211 VAATLLVABEALGFTPEAE-GLSAPLSHCCFCRLAFRNLGALLPRHLQCCMRPR 269
Db 365 -AASALLASTSAIWAAPAPEAKAAAGAEITWAYELGAGLGAIFGLLSLSSFSASIRLPA 423
QY 270 TLRLFLVACLCSWVALMTFTLFYDFVGEV--LYQGVV--RAEPTGEARRH---YDEGVR 322
Db 424 GLEAQETARASSM-----GEAVQLANSUPPTQGOAILDAARHAFIWSHVA 470
QY 323 MGSIG-LFLOCAISLVFSL 340
Db 471 LSSAGSMILLAVGMWFSL 489
RESULT 15
Y051_MYCTU STANDARD; PRT; 640 AA.
AC Q10880; O53628;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein RV0083/MT0090.
GN RV0083 OR MT0090 OR M1C251.01 OR MV030.27.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
ON NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:22:38 ; Search time 149.566 Seconds
(without alignments)
1166.587 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861
Sequence: 1 MVQRLWVSLRLHRRKAQLL.....AIYFATQVVFDSKLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP_TREMBL 25:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_phase:*
- 11: sp_plant:*
- 12: sp_rodent:*
- 13: sp_virus:*
- 14: sp_vertebrate:*
- 15: sp_unclassified:*
- 16: sp_rvirus:*
- 17: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	4 Q96JT2	Q96jt2 homo sapien
2	2798	97.8	553	6 Q95X15	Q95x15 macaca fasc
3	2602	90.9	553	11 Q8X0H7	Q8X0h7 mus musculus
4	2531	88.5	501	6 Q95KC5	Q95kc5 macaca fasc
5	2099	73.4	450	11 Q8X252	Q8x252 mus musculus
6	519.5	18.2	599	5 Q9VSV1	Q9vsv1 drosophila
7	512	17.9	112	11 Q8R1I0	Q8r1i0 mus musculus
8	494.5	16.9	530	11 Q8C204	Q8c204 mus musculus
9	456	15.9	576	13 Q8C274	Q8c274 oryzias lat
10	446	15.6	751	11 Q8X4S3	Q8x4s3 rattus norv
11	394.5	13.8	754	5 Q869P1	Q869p1 dictyostel
12	347.5	12.1	515	10 Q65803	Q65803 daucus caro
13	347.5	12.1	501	10 Q9FNR6	Q9fnr6 daucus caro
14	346.5	12.1	501	10 Q9SQK6	Q9sqk6 vitis vinif
15	346.5	12.1	515	10 Q9SLN7	Q9sln7 daucus caro
16	346	12.1	512	10 Q9SQK5	Q9sqk5 vitis vinif

17	344.5	12.0	520	10 Q7XA53	Q7xa53 glycine max
18	338	11.8	512	10 Q9S7Z5	Q9s7z5 apium grave
19	337.5	11.8	523	10 Q04077	Q04077 vicia faba
20	336.5	11.8	501	10 Q9SP63	Q9sp63 vitis vinif
21	333.5	11.7	599	10 Q7X9X8	Q7x9x8 plantago ma
22	333	11.6	512	10 Q9ZIB9	Q9zib9 apium grave
23	332.5	11.6	508	10 Q8VYX3	Q8vyx3 brassica ol
24	331	11.6	511	10 Q84RQ3	Q84rq3 lotus japon
25	330	11.5	516	10 Q43653	Q43653 solarum tub
26	329.5	11.5	524	10 Q9XHL6	Q9xhl6 pisum sativ
27	328.5	11.5	506	10 Q944W2	Q944w2 oryza sativ
28	328.5	11.5	594	10 Q80605	Q80605 arabidopsis
29	327.5	11.4	506	10 Q948L0	Q948l0 oryza sativ
30	325.5	11.4	604	10 Q9FVL6	Q9fvl6 lycopersico
31	324.5	11.3	502	10 Q9SP14	Q9sp14 alonsoa mer
32	324.5	11.3	512	10 Q8RWQ6	Q8rwq6 arabidopsis
33	323.5	11.3	512	10 Q39231	Q39231 arabidopsis
34	323.5	11.3	512	10 Q80550	Q80550 arabidopsis
35	323.5	11.3	595	10 Q84KR4	Q84kr4 oryza sativ
36	317.5	11.1	507	10 Q40583	Q40583 nicotiana t
37	317.5	11.1	530	10 Q9M535	Q9m535 euphorbia e
38	316.5	11.1	500	10 Q9FVJ6	Q9fvj6 lycopersico
39	313.5	11.0	605	10 Q7Y1Q0	Q7y1q0 solarum tub
40	313	10.9	521	10 Q9SXM0	Q9sxm0 zea mays (m
41	311	10.9	533	10 Q41152	Q41152 ricinus com
42	309.5	10.8	607	10 Q8LPM4	Q8lpm4 citrus sine
43	307.5	10.7	488	10 Q9FV92	Q9fv92 solarum tub
44	306.5	10.7	513	10 Q8VYX4	Q8vyx4 brassica ol
45	306.5	10.7	528	10 Q8LPM6	Q8lpm6 citrus sine

ALIGNMENTS

RESULT 1

Q96JT2	PRELIMINARY;	PRT;	553 AA.
ID	Q96JT2		
AC	Q96JT2;		
DT	01-DEC-2001 (TRENBLrel. 19, Created)		
DT	01-DEC-2001 (TRENBLrel. 19, Last sequence update)		
DT	01-JUN-2003 (TRENBLrel. 24, Last annotation update)		
DE	Protein.		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Prostate;		
RX	MEDLINE=21139094; PubMed=11245466;		
RA	Xu J., Kalos M., Stolk J.A., Zasloff E.J., Zhang X., Houghton R.L.,		
RA	Filho A.M., Nolasco M., Badaro R., Reed S.G.;		
RT	"Identification and characterization of prostate, a novel prostate-		
RT	specific protein.";		
RL	Cancer Res. 61:1563-1568(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Peripheral Nervous System;		
RX	MEDLINE=23388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.W., Bhat N.K.,		
RA	Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peripheral Nervous System;
 RA Strausberg R.;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY033593; AAK54386.1; -;
 DR EMBL; BC050416; AAH50416.1; -;
 SQ SEQUENCE 553 AA; 59322 MW; 0AF23FBC742A667 CRC64;

Query Match 100.0%; Score 2861; DB 4; Length 553;
 Best Local Similarity 100.0%; Pred. No. 3.4e-210;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLRLHRKKAQLLLVNLITFGLEVCLAAGITYVPPILLVGVVEEKFTWVLGIG 60
 DB 1 MVQLWVSRLRLHRKKAQLLLVNLITFGLEVCLAAGITYVPPILLVGVVEEKFTWVLGIG 60
 QY 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSGLILLSLFIPIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSGLILLSLFIPIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVLLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVLLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDMTSALAPYLGTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 DB 181 IDMTSALAPYLGTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 QY 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
 DB 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
 QY 301 YQGVPAEPCTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPAEPCTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAIFYATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAIFYATQ 540
 QY 541 VVFDKSLAKYSA 553
 DB 541 VVFDKSLAKYSA 553

RESULT 2

Q95K15 PRELIMINARY; PRT; 553 AA.
 AC Q95K15;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;

RN SEQUENCE FROM N.A.
 RP TISSUE=Temporal cortex;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB060851; BAB46871.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

Query Match 97.8%; Score 2798; DB 6; Length 553;
 Best Local Similarity 98.2%; Pred. No. 2.2e-205;
 Matches 542; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVQLWVSRLRLHRKKAQLLLVNLITFGLEVCLAAGITYVPPILLVGVVEEKFTWVLGIG 60
 DB 1 MVQLWVSRLRLHRKKAQLLLVNLITFGLEVCLAAGITYVPPILLVGVVEEKFTWVLGIG 60
 QY 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSGLILLSLFIPIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSGLILLSLFIPIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVLLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVLLDFCGQVCTPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDMTSALAPYLGTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 DB 181 IDMTSALAPYLGTQBECLFGLLTFLITCTVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 QY 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
 DB 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
 QY 301 YQGVPAEPCTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPAEPCTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAIFYATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAIFYATQ 540
 QY 541 VVFDKSLAKYSA 552
 DB 541 VVFDKSLAKYSA 552

RESULT 3

Q8K0H7 PRELIMINARY; PRT; 553 AA.
 AC Q8K0H7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DE Hypothetical protein (PROSTEIN homolog).
 GN PROSTEIN OR 2210413P12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

```
RA Strausberg R.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
DR EMBL; BC031381; AAH31381.1; -.
DR MBL; AK035428; BAC29063.1; -.
DR MGD; MGI:1922082; Prostein.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0045285; C:biquinol-cytochrome-c reductase complex; IEA.
DR GO; GO:0008121; F:biquinol-cytochrome-c reductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005805; Rieske.
DR PROSITE; PS00200; RIESKE_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59742 MW; 6A474FEA630B4B13 CRC64;

Query Match      30.9%; Score 2602; DB 11; Length 553;
Best Local Similarity 90.9%; Pred. No. 2e-190;
Matches 502; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLVNLLTFGLEVCVLAAGTYVPPILLEVGEVEKEWTWVLGIG 60
Db 1 MIQRLWASRLRRHKAQLLVNLLTFGLEVCVLAAGTYVPPILLEVGEVEKEWTWVLGIG 60

QY 61 PVGLGVCPVLLGSASDHWRGRRRPFVWLSGLSLLSLFIPRAGWLAGLCPDPRPL 120
Db 61 PVGLGVCPVLLGSASDHWRGRRRPFVWLSGLSLLSLFIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVLLDFCGQVCTPLEALLSLDFRDPDHCROAYSVYAFMISLGGICGLVLLPA 180
Db 121 ELALLILGVLLDFCGQVCTPLEALLSLDFRDPDHCROAYSVYAFMISLGGICGLVLLPA 180

QY 181 IDWDTSAAPYLGTOECLFGLLTFLTCVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240
Db 181 IDWDTSAAPYLGTOECLFGLLTFLTCVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240

QY 241 CCPCARLARLAFNLGALLPRLHQLCCMPETLRLFVAELCSNMALMTFTLFTDFVGEGL 300
Db 241 CCPCARLARLAFNLGALLPRLHQLCCMPETLRLFVAELCSNMALMTFTLFTDFVGEGL 300

QY 301 YQGVPAEPTGEARRHYDEGVRMSGLGLFLQCAISLVFSLVMDRLVQREGTRAVYLASVA 360
Db 301 YQGVPAEPTGEARRHYDEGVRMSGLGLFLQCAISLVFSLVMDRLVQREGTRAVYLASVA 360

QY 361 APVAAAGATCLSHVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAAGATCLSHVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGSGGLPPPPALCGASACDVSVRVVVEGPTFA 480
Db 421 ASSEDSLMTSFLPGKPGAPFPNGHVAGSGGLPPPPALCGASACDVSVRVVVEGPTFA 480

QY 481 RVVPGIGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGVVAYFATQ 540
Db 481 RVVPGIGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGVVAYFATQ 540

QY 541 VVFDKSDLAKYS 552
Db 541 VVFDKNDLAKYS 552

RESULT 4
Q95KC5
ID Q95KC5
AC Q95KC5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)

PRELIMINARY; PRT; 501 AA.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

01-DEC-2001 (TReMBLrel. 19, Last sequence update)
01-OCT-2003 (TReMBLrel. 25, Last annotation update)
Hypothetical protein.
Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecidae; Macaca.
NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062977; BAB60745.1; -.
KW Hypothetical protein.
SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04BE0470 CRC64;

Query Match      88.5%; Score 2531; DB 6; Length 501;
Best Local Similarity 97.8%; Pred. No. 4.8e-185;
Matches 489; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 53 MTWVIGIPVGLVGVCPVLLGSASDHWRGRRRPFVWLSGLSLLSLFIPRAGWLAGL 112
Db 1 MTWVIGIPVGLVGVCPVLLGSASDHWRGRRRPFVWLSGLSLLSLFIPRAGWLAGL 60

QY 113 LCPDPRPLEALLILGVLLDFCGQVCTPLEALLSLDFRDPDHCROAYSVYAFMISLGG 172
Db 61 LCPDPRPLEALLILGVLLDFCGQVCTPLEALLSLDFRDPDHCROAYSVYAFMISLGG 120

QY 173 CLGYLLPAIDWDTSAAPYLGTOECLFGLLTFLTCVAATLLVAEEAALGPTEPAEGL 232
Db 121 CLGYLLPAIDWDTSAAPYLGTOECLFGLLTFLTCVAATLLVAEEAALGPTEPAEGL 180

QY 233 SAPLSLPHCCPCARLARLAFNLGALLPRLHQLCCMPETLRLFVAELCSNMALMTFTLFTY 292
Db 181 SAPLSLPHCCPCARLARLAFNLGALLPRLHQLCCMPETLRLFVAELCSNMALMTFTLFTY 240

QY 293 TDFVGEGLYQGVPAEPTGEARRHYDEGVRMSGLGLFLQCAISLVFSLVMDRLVQREGTR 352
Db 241 TDFVGEGLYQGVPAEPTGEARRHYDEGVRMSGLGLFLQCAISLVFSLVMDRLVQREGTR 300

QY 353 AVYLASVAAFPVAAAGATCLSHVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLP 412
Db 301 AVYLASVAAFPVAAAGATCLSHVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLP 360

QY 413 KYRGDTGASSEDLSMTSFLPGKPGAPFPNGHVAGSGGLPPPPALCGASACDVSVRV 472
Db 361 KYRGDTGASSEDLSMTSFLPGKPGAPFPNGHVAGSGGLPPPPALCGASACDVSVRV 420

QY 473 VVGSGPTGARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGL 532
Db 421 VVGSGPTGARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGL 480

QY 533 VAIYFATQVVPDKSDLAKYS 552
Db 481 VAIYFATQVVPDKSDLAKYS 500

RESULT 5
Q95KC5
ID Q95KC5
AC Q95KC5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
Hypothetical protein (Fragment).
PROSTEIN OR 2210413P12R1X.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```

OX NCB1_TaxID=10090;
RN [1]
RA SEQUENCE FROM N.A.
RP Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC034084; AAH34084.1; -
DR MGD; MG1:1922082; Protein.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0045285; C:ubiquinol-cytochrome-c reductase complex; IEA.
DR GO; GO:0008121; F:ubiquinol-cytochrome-c reductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005805; Rieske.
DR PROSITE; PS00200; RIESKE_2; 1.
KW Hypothetical protein.
FT NON_TER
SQ
  Query Match      73.4%; Score 2099; DB 11; Length 450;
  Best Local Similarity 90.0%; Pred. No. 4.1e-152;
  Matches 404; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

QY 104 PRAGWAGLCCPPRPLEALILGVLLDFCGVCFPLEALLSLDFRDPDCHCRQAYSV 163
Db 1 PRAGWAGLCCPPRPLEALILGVLLDFCGVCFPLEALLSLDFRDPDCHCRQAYSV 60

QY 164 YAFMISLGGCLGYLLPAIDWDTLSALAPYLGTQBECLFGLTLTIFLTCVAATLLVAEAAAL 223
Db 61 YAFMISLGGCLGYLLPAIDWDTLSALAPYLGTQBECLFGLTLTIFLTCVAATLLVAEAAAL 120

QY 224 GPTEPAGLSAPSLCHPCPCRLARLAFNLGALLPRHLQCCMPRTLRLRFLVAELCSWM 283
Db 121 GPTEPAGLSAPSLCHPCPCRLARLAFNLGALLPRHLQCCMPRTLRLRFLVAELCSWM 180

QY 284 ALMTFTFLYDFVGEGLYQVPRABGTARRHYDGRVMSGLGLFLOCAISLVFSLVMD 343
Db 181 ALMTFTFLYDFVGEGLYQVPRABGTARRHYDGRVMSGLGLFLOCAISLVFSLVMD 240

QY 344 RLVOREGTRAVYLASVAAPFPAAGATCLSHSVAVVTASAAALTCFTFSAALQILPYTLASLY 403
Db 241 RLVOREGTRAVYLASVAAPFPAAGATCLSHSVAVVTASAAALTCFTFSAALQILPYTLASLY 300

QY 404 HREKQVFLPKYRGDTGASSEDLSMTSLFPGPKGAPPFNHGVAGSGGLPPLPPALCCA 463
Db 301 HREKQVFLPKYRGDTGASSEDLSMTSLFPGPKGAPPFNHGVAGSGGLPPLPPALCCA 360

QY 464 SACDVSRRVVVGGPTTARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSHSVTAY 523
Db 361 SACDVSRRVVVGGPTTARVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSHSVTAY 420

QY 524 MVSAAAGLGLVAYIFATQVVPFDSKIAKYS 552
Db 421 MVSAAAGLGLVAYIFATQVVPFDSKIAKYS 449

RESULT 6
QYVSU1 PRELIMINARY; PRT; 599 AA.
AC QYVSU1
DT 01-MAY-2000 (T:EMBLrel. 13, Created)
DT 01-MAY-2000 (T:EMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (T:EMBLrel. 14, Last annotation update)
DE CG4484 protein.
DE CG4484.
GN CG4484.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCB1_TaxID=7227;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```


Query Match 15.98; Score 456; DB 13; Length 576;
 Best Local Similarity 24.04; Pred. No. 1.5e-26;
 Matches 144; Conservative 95; Mismatches 198; Indels 164; Gaps 17;

QY 12 RHRKAQALLVNLITFTGLEVCCLAAAGITVPPLLLEVGVEEKFMTWLVGIGPVGLVLCVPL 71
 DB 56 RRSRGRLLHSMWFGREFCYAVEAFVTVLLSVGLPRSLVSLWLSPIGLFLQPII 115
 QY 72 GSASDHWGRYGRRRPIWALSIGIL-----LSLPL-----IPRAGWLAGLLC 114
 DB 116 GSASDYCRSSWGRRRPYI--LVLGILMLVGLSFLNGDAVWSLVSDRSRSTW----- 167
 QY 115 PDPRLELALLVGLGLDPCGQVCTPLEALLSDL--FRDPDHCQAYSVVAFMISLGG 172
 DB 168 -----AIVVMGVVDFDAADFDGPIKAYLFDVCSYQDKB--RGLHYHALFTGLGG 218
 QY 173 CLGYLLPAIDWTSALAPYLGTQBECLF-----GLLTILIFITCV----- 211
 DB 219 ACGVLGAMDWGSHVILGRLLGSEYQVIYFVSALTWGVFLIVHLFSIPEKPLAKVSESSA 278
 QY 212 -AATLVAEEA---ALG--PTPEAGLSAPSLSPHCCPC----- 244
 DB 279 SSALRLLGHSNGYGALGKEFVSPVPTSPFIRPSYSALGRPRPSFALGEANSVTSS 338
 QY 245 -----RARLAFRNLGALLPRHLQCCRMPTLRRLFAELCSWMALMTFTLFYTFDV 296
 DB 339 AKOPIKEDQKMTFRS-----LMKAIENMNPVHYFLCISHLLGWAAFLCNMLPFTDFM 391
 QY 297 GEGLYQGVPAEPTGTRARRHYDGVNMGSLGLFLQCAISLVFSLYMDRLVQRFTRAVYL 356
 DB 392 GOIVYRGNPYAENSTAYITIERGVGVGCGICINAVSALYSYVQRFLLPYIGLKGLYF 451
 QY 357 ASVAAPFAAGACGLSHSVAVATSAALTGFTFSALQILPYTLASLYHREKOVFPKYRG 416
 DB 452 MGTVFVGMGTSLLIGLPEVIATILCSVFGVMSSTLYTIPFNLIAYQEEBEEQV-KLEG 510
 QY 417 DTGASSEDLSMTSFUPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVGE 476
 DB 511 -----GNESPRGTGM-----DCAALTVCVQL----- 531
 QY 477 PTBAAVVGRCIGICLDLALDLSAFLLSQVAPSLFMSGIVOLSQVTAVMYSAAGLGLVAIY 536
 DB 532 ---AQIIVGAG-----LGLAVNAAGSVIVVLSASSISLIGCI 566
 QY 537 F 537
 DB 567 F 567

RESULT 10
 Q8K453 ID Q8K453 PRELIMINARY; PRT; 751 AA.
 AC Q8K453;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Proton-associated sugar transporter A.
 GN PAST-A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Shimokawa N., Okada J., Haglund K., Dikic I., Koibuchi N., Miura M.;
 RT "Past-A, a new class proton-associated sugar transport facilitator in
 RL the central nervous system.";
 DR EMBL; AB075229; BAB97313.1; -
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 KW Sugar transport.

SQ SEQUENCE 751 AA; 81751 MW; 5C33E1C2E4A1B2EC CRC64;
 Query Match 15.63; Score 446; DB 11; Length 751;
 Best Local Similarity 22.33; Pred. No. 1.2e-25;
 Matches 160; Conservative 79; Mismatches 206; Indels 272; Gaps 15;

QY 17 QLLVNLITFTGLEVCCLAAAGITVPPLLLEVGVEEKFMTWLVGIGPVGLVLCVPLGSASD 76
 DB 86 ELLFNGCILGIEFSYAMETAYTVPLQNGLPOLYSIVWFISPIGLFLQLLQAWS 145
 QY 77 HWRGRRRRPFTWALSIGILLSFLIPRAGWLAGLLCPDP---RPLELALLILVGLLD 133
 DB 146 RCTSRGRRRRPFIIVLAIGALLGLSLLN--GRDIGHALADTATNHNKIGILLTCGVVLM 204
 QY 134 FCGQVCTPLEALLSDLFRDPDHCQAYSVVAFMISLGGCLGYLLPAIDWTSALAPYL 193
 DB 205 FSASADNPSHAYMDCVGPVDQR--GLNTHALMAGLGGFGVVGIIHWDKTSFGRALG 263
 QY 194 TQBECLFGLLTILFELTCAATLLVA-----CRALAFENLGA--EAAA 222
 DB 264 GOLRAVIV-IFTAITLSTVTTVLVSIPEPLRPLGKRTAMKSPSLPLPPSPVLLBEA 322
 QY 223 LGPTPE---AEGLSAPSLSPHCCP----- 243
 DB 323 -GDTLPSTTATSLYASPSPIPPSPPLTPKYGFISRDSSLTGINFASFGTSNIDSVL 381
 QY 244 -----CRALAFENLGA-- 257
 DB 382 IDCFTAGHDNYLALPSSVPRQAISVSFPRAPDGFYCOERGLERREGPLTLGLDGLVRG 441
 QY 258 -----PR----- 259
 DB 442 SLDTSKPRASGILKRPQTLALPDVAGNGPETSRRNRVTFSSQOVANILLNGVYSELTG 501
 QY 260 -----LHQLCC---EMPRTLRLFAELCSWMALMTFTLFYTFDFVGEGLYQGVPR 306
 DB 502 SSEQSEQLSLRLCSTIYNPVRNLCVNHFLGWSFEGMLLFTYDFMGEVVFQGDPK 561
 QY 307 AEPCTEARRHYDEGVNMGSLGLFLQCAISLVFSLYMDRLVQRFTRAVYLASVAAPFAA 366
 DB 562 APHASEAYKYNQSGVTMGCMCIYAFSAAFYSAILEKECLSVRTLFIAYLLFQLGT 621
 QY 367 GATCLSHSVAVATSAALTGFTFSALQILPYTLASLYHREKOVFLPKYRGDTGASSED 426
 DB 622 GLATSLRNLYVLSLCTHYGILFTLCTLPSYLLCDYIYQSK---KFAG-----SSADG 671
 QY 427 LMTSFLPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGR 486
 DB 672 TR-----R 674
 QY 487 GICLDLALDLSAFLLSQVAPSLFMSGIVOLSQVTAVMYSAAGLGLVAIYFATQVVF 543
 DB 675 GMGVDISLLSCQVFLAQILVSLVIGP-----LTSAVGSANGVMYFASLVSF 720

RESULT 11
 Q869P1 ID Q869P1 PRELIMINARY; PRT; 754 AA.
 AC Q869P1;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to Arabidopsis thaliana (Mouse-ear cross). sucrose transporter
 DE protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910;
 RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
 Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AC117072; A052628.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005828; Sub transporter.
 DR Pfam; PF00083; sugar tr; 1.
 SQ SEQUENCE 754 AA; 82774 MW; 614C473C832ACCE4 CRC64;
 Query Match 13.8%; Score 394.5; DB 5; Length 754;
 Best Local Similarity 23.5%; Pred. No. 1e-21;
 Matches 128; Conservative 90; Mismatches 200; Indels 127; Gaps 13;
 QY 13 HRKA---OQLLVNLTFLGLEVCLAAAGITYVPPLLLEVGVEKFMVTLGIGPVLGLVCV 68
 DB 197 HKTLPLFNLCITICFLGVQFGWALQIAFSTPLFLELGVQKWSYIWLAPISGLIVQ 256
 QY 69 PLGASDHRGGRYGRRRPFIWALSGLLSLFLIPRAGWLAGLL--CPDPRLELALI 126
 DB 257 PLVGVTIDRSECRFGRKPFILIGSVFISIGLVLSNAETFGSYFGDSQKKSIAISAI 316
 QY 127 LGVGLDFGQVCFTEALLSLDFPDHCRQAYSVYAFMISLGCGLGLYLPDAIDWTS 186
 DB 317 VGVWILDLSNNAVOACRALLVDI--AAPSOOSLGSLFSLMLGTGNLLGYMMSI--DLV 373
 QY 187 ALAPYLGTOBECLFGLLTFLFTCVAAATL--LVAAEAAAGTFFAEGLSAPSLSPHCCPC 245
 DB 374 RMVPEMKTDTRALFTLSIMVLLFCVVMTLGFVTEEQYIRVNEQ--QSVENP----- 422
 QY 246 ARLAFNLGALLPRLHQLCCMRPTLRRLFVABLCSWMAIMTFLTYTDFVGEGLYQVGP 305
 DB 423 -----LKTMPKIVKMPFTYLQRLCAVQFFSWGMFVLFTTIVGVNVPFGSDP 471
 QY 306 RAEPGTAEARRHYDEGVKMSGLFLQCAISLVFSLVMDLVQFRTGTRAVYLASVAAPVA 365
 DB 472 NAEYSDSRLTFQDGVKWSLSLTSSGITIAVSLIPVLKFDIMKYIYIGNL----- 526
 QY 366 AGATCLSHSVAVVTAS-----AALTGTFSAQLPILPYTLASLYHREKQVLPKYRG 418
 DB 527 --LQCFIFALFYVESKISGLLLIAGTGPWATVMLPESIVGM----- 568
 QY 419 GGASSEDLSMTSLPGKPGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEP 478
 DB 569 -GVEDNES-----SG----- 577
 QY 479 EARVPEGRGICLDLAIDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVAAGLGLVAIXPA 538
 DB 578 -----LNIQTINIFVVVPQMVSLGIGLILDSKGNVYSLTA--GSVASPFA 623
 QY 539 TQVVF 543
 DB 624 TLFCF 628
 RESULT 12
 O65903
 ID O65803 PRELIMINARY; PRT; 515 AA.
 AC O65803;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Sucrose/H+ symporter.
 GN SUT2.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
 OC Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nantaise; TISSUE=Root;
 RX MEDLINE=99063785; PubMed=9847123;
 RA Shakya R., Sturm A.;
 RT "Characterization of source- and sink-specific sucrose/H+ symporters
 from carrot";
 RL Plant Physiol. 118:1473-1480(1998).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 DR ENBL; Y16768; CAA76369.1; -
 DR PIR; T14340; T14340.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008515; F:sucrose transporter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0015770; P:sucrose transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005989; Suc7H_symport.
 DR Pfam; PF00083; sugar tr; 1.
 DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
 KW Transmembrane.
 SQ SEQUENCE 515 AA; 54424 MW; 897B90657C9E243C CRC64;
 Query Match 12.1%; Score 347.5; DB 10; Length 515;
 Best Local Similarity 25.9%; Pred. No. 2.6e-18;
 Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
 QY 17 QLLVNLTLFGLVCLAAAGITYVPPLLLEVGVEKFMVTLGIGPVLGLVCVPLLGASD 76
 DB 34 KLVVAAIAAGVQFGWALQSLTLPYVQLLIGPHKAAIWLQGPISGMLVQPTVGYSD 93
 QY 77 HWRGRYGRRRPFIWALSGLLSLFLIPRAGWLAGL-----LCPDPRLELALLILGV 129
 DB 94 HCQSFGRRRPFIAGAGCAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGP 150
 QY 130 GLLDFGQVCFTEALLSLDFR--DPDHCQAYSVYAFMISLGCGLG-----LL 178
 DB 151 WILDVANNMLQPCRALLDLCSGDTRRMSANAFYFFMAVGNILGYAAGSYNNLYKLF 210
 QY 179 PAIDWDTSAAPYLGTOECLFGLTLFLTCVAAATLLVAEEAALGTEPAEGLSAPSL 238
 DB 211 PFS--KTHACDLYCANLKSCFIISIALIIITVVALSVRENS--GPPDDAABEEP--- 263
 QY 239 PHCCPCRARLAFNLGALLPRLHQLCCMRPTLRRLFVABLCSWMAIMTFLTYTDFVGE 298
 DB 264 ----PSSGKIYV--FGELLGALKD-----FAPMLLLIIVTCLNWIAPFPLFTDWMGR 313
 QY 299 GLYGVPRPAEPGTAEARRHYDEGVKMSGLFLQCAISLVFSLVMDLVQFRTGTRAVYAS 358
 DB 314 EIVGTT--AGQG----KLYDQGVRAAGLGLLNSVWLGLTSAIVEYLVRGVGVKI--LWG 366
 QY 359 VAAPFVAAGATCL-----SHSV-----AVVTASA---ALTGTFSAQLIL 395
 DB 367 FVNFIILGLVMTVVVSKVAHQHREHANGQLLPSPAGVKAGALSLSILGIPLSIYSI 426
 QY 396 PYTLASLYHREKQVFLPKYRGDTGCASSEDLSMTSLFPGPKPG-----APFPNGHVGA 448
 DB 427 PFALASIVSSGSGAGQGLSLGLVNLAIWVPMIVSVLAGPFDLSFGGNLPAFVVGAISA 486
 QY 449 GSGSLP-----PPPALCGASACDVS 469
 DB 487 AISGVLAIVLPLPKPSKAAASKLSLS 511
 RESULT 13
 O3FNR6
 ID Q3FNR6 PRELIMINARY; PRT; 515 AA.
 AC Q3FNR6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```

01-MAR-2001 (TREMBlrel. 16, Last sequence update)
01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SUCROSE/proton symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Apiales; Apiaceae; Scandiceae; Daucinae;
OC Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Sturm A.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Shkya R.;
RL Thesis (2000), Department of Botany,
RL University of Basel, Basel, Switzerland.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AJ303199; CAC19689.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008515; F:sucrose transporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0015770; P:sucrose transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005989; Suc/H symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54425 MW; 9580978C726523DC CRC64;

Query Match 12.1%; Score 347.5; DB 10; Length 515;
Best Local Similarity 25.9%; Pred. No. 2.6e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;

QY 17 QLLVNLTLFGLVCLAAAGITYPPPLLELVGVEEKMFTWVLGIPVLGVCVPLLGASD 76
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 34 KLVVAAIAAGVGFQWALQSLTLTPYVQLLGIPHKWAAVILWCGPISGMLVQPIVGYSD 93

QY 77 HWRGVRRRPPTWALSGLLGLSLFLIPRAGWLAGL-----LCPDPRLEALLILGV 129
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 94 HCQSFGRRPPTIAGAGCVAISVILI---GFAADISYKAGDDMSKTLKPRATVTVFVIGF 150

QY 130 GLIDFCGVCFTEPLEALLSLR-DFDHCROAVSVVAFMISLGGCLGY-----LL 178
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 WILDVANNLQGCRCALLADLCSGDFTRRWSANAFYSFFMAVGNILGYAAGSYNNLYKLF 210

QY 179 PAIDWDTALAPYLGTCBCLFGLLTLIFLTCTVAATLLVAEEAALGPTPEAGLSAPLS 238
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 211 PFS--KTHACDLKCANLCKSFIISALLIITVVALSVVRENS--GPPDDADAEEP--- 263

QY 239 PHCCPCBARLAFNLGALLPRHLQCCMRPTRLRLFVAELCSWALMTFTLYFTDFVGE 298
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 264 ----PSSGKIPV--FGEILGALKDL----PRMLLLIIVTCNLNIAWPFILFDTDWGR 313

QY 299 GLYQGVPRABPGTEARRHYDEGVGMGSLGLFLQCAISLVSLVMDRLVORFTRAVILAS 358
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 314 EIYGGT--AGKG---KLYDQGVACALGULLNSVVLGTSIAVEYLVRGVGVKI-LWG 366

QY 359 VAAPFVAAGATCL-----SHV-----AVVTASA----ALTGTFPSALQIL 395
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 367 FVNFIILAIGLVMTVVVKVAQHOREHSANGQLLPPSAGVKAGALSLSILGIPLSITYSI 426

QY 396 PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGKPG-----APFPNGHVGA 448
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 427 PFALASISYSSGGAGGGLSGLVNLALVVPQMTIVSVLAGPFDLSFGGGLNLPFVVGAI 486

QY 449 GSGLLP-----PPALCGASACDVS 469

01-MAR-2001 (TREMBlrel. 16, Last sequence update)
01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SUCROSE/proton symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Apiales; Apiaceae; Scandiceae; Daucinae;
OC Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Sturm A.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Shkya R.;
RL Thesis (2000), Department of Botany,
RL University of Basel, Basel, Switzerland.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AJ303199; CAC19689.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008515; F:sucrose transporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0015770; P:sucrose transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005989; Suc/H symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 501 AA; 53938 MW; 4D0B4DE2EF2F4BA8 CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 501;
Best Local Similarity 27.0%; Pred. No. 3e-18;
Matches 153; Conservative 83; Mismatches 197; Indels 133; Gaps 22;

QY 17 QLLVNLTLFGLVCLAAAGITYPPPLLELVGVEEKMFTWVLGIPVLGVCVPLLGASD 76
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 29 RLRVASVACGIQFWALQSLTLTPYVQELGIPHAWSSIIWLCGPLSGLLVQVLVGLSD 88

QY 77 HWRGVRRRPPTWALSGLLGLSLFLIPRAGWLAGL--CPDPRLEALLILGVGLDF 134
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 89 RCNSRFGRRPPIVAGATSIWVAVLIIGFSADIGLLGDCADRPRPRAVATFVVGFWLLDV 148

QY 135 CGQVCFPLPALLSDLRDPDH--CROAVSVYAFMISLGGCLGYLPA-----INDTS 186
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 ANNVQGPCRALLADL--TEKDHRRTRVANAYFSLFIAVGNVLGATGYSYSGNFRIFWFTS 207

QY 187 ALAPYLGTQCECL---FGLTLIFLTCTVAATLLVAEEAALGPTPEAGLSAPLSPHCC 242
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 208 TSS-----CNADCANLKSFAFLDIIF--IAITYISITAA-----QELPLSSSRSTHIS 255

QY 243 -----PCRAELAFNLGALLPRHLQCCMRPTRLR-----LFVAELCSWALMTFT 289
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 256 EEMAEASTHAQEAFF-----LWELF---GTLRYLSGSIWILFVTAL-TWIGLLPFL 301

QY 290 LFTDFVGEGLYQGVPRABPGTEARRHYDEGVGMGSLGLFLQCAISLVSLVMDRLVORF 349
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 302 LFTDFDMGRIRYIGKPK-----NEGQNYNTGVNMGALGMLNSVVLGITSVLMEKLCRKW 355

QY 350 GTRAVILASVAAPFVAAGATCLSHSVAVVTASAALTGTFPSALQILPYTLASLYHREKQV 409
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 356 GAGFVW-----GLSNILMSLCFLMLILS---AVVKEMD--- 386

QY 410 FLPKYRGDTGGASSEDLSMTSFLPGKPGAPFNGHVAGGSGLLPPFALCGASACDVS 469

```



```
Db 387 -----FLGHDL-----PSGVIA-----ALIVFSILGIP 411
Qy 470 VRVWGP---TEARVP---GRGICLDLAILDSAFILSOVAPSLFMGSIVOL-----SOSV 520
Db 412 LAITYVPALSTRIESLGLGGLSMGVNLAIVIPQIVSLGSGPWDLFGGNSPSL 471
Qy 521 TAYWSAAGLGLVAIFYATQVFDKS 546
Db 472 AVAAVAFAAGLVAIILAIIPRSSADKS 497

RESULT 15
Q9SLN7 PRELIMINARY; PRT; 515 AA.
AC Q9SLN7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sucrose transporter protein.
GN CSUT.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Apiales; Apiaceae; Apiaceae; Scandiceae; Daucinae;
OC Daucus.
OC NCBI_taxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Wu N., Zhao F., Zhang L., Huang M.;
RT "Sucrose transporter protein."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AB036758; BAA89458.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008515; F:sucrose transporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0015770; P:sucrose transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005989; Suc/H_symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane
SQ SEQUENCE 515 AA; 54523 MW; 2B548D1D9DFF51AC CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 515;
Best Local Similarity 27.2%; Pred. No. 3.1e-18;
Matches 137; Conservative 76; Mismatches 215; Indels 75; Gaps 19;

Qy 17 QLLIVNLTFTGLEVCLAAGITYVPPLLLEVGVEEKEFTWVLGIPVGLVCVPLIGSASD 76
Db 34 KLVIVAAIAGVQFGWALQSLTTPYVQLLGIPIKRWAAIWLCEPISGMLVQPIVGYSD 93
Qy 77 HWRGRRRRFFIWAISLGLSLFLIPRA---GWLAG-LLCPDRPLEIALLILGVGLL 132
Db 94 HCOSFGRRRFFIASGAGCAISVILLIGPAADIGYKAGDDMSKTLKPRAVTVFVIGFWIL 153
Qy 133 DFCGVCTFTLEALLSDLFR-DPDHCRQAYSVYAFMISLGGCLGY-----LLPAL 181
Db 154 DVANNMLQGFCEALLADLCSGDTRMRKSANAFYFFMAVGNILGYAAGSYNNLYKLFPPS 213
Qy 182 DWDTSAALPYLGTQEEC-LFGLLTITLTCTVAATLLVAEEAALGPTPEAGISAPLSPH 240
Db 214 --KTHACDLYCANLKSFIISIALIIITVA--LSVREKQWSP-DDADADEP----- 263
Qy 241 CCPCRARLAFNIGALLPRHLQCCNRPVRLRLFVAELCSNMALMTFTLYTDFVGEGL 300
Db 264 --PSSGKIPV--FGEILGALKDL-----PRPMLLLIIVTCLNWIATWFFIIFLFDWMGREI 315
Qy 301 YQGVPRAPGTEARRHHYDEGVKMSLGLFLQCAISLVSFLVMDRLVQRFCRAVILASVA 360
```

```
Db 316 YGGT--AGKG---KLYDQGVRAAGSLGLLLNSVVLGUTSTAVEYLVKVGGVKI-LMGVV 368
Qy 361 APPVAAGATCL-----SHSV-----AVVTASA-----ALTGFTFSALQILPY 397
Db 369 NFILAIGLVMTVVVSKVAQHREHSANGQLLPPSAGVKAGALSLSILGIPLSITYSIPIF 428
Qy 398 TLASLYHREKQVFLPKYRGDTGGASSEDLSMTSLPQPKFG-----APPNCHVGAGG 450
Db 429 ALASIYSSGGAGQGLSLGVNLNLAIVVPQMTIVSVLAGPFDLSLFGGGLNLPAPVVGAIISAAI 488
Qy 451 SGLLP---PPALCGASACDVS 469
Db 489 SGVLAIVLLPKPKDAASKLSLS 511
```

Search completed: August 25, 2004, 02:31:18
Job time : 154.566 secs

This Page Blank (uspto)